REPORT RESUMES

ED 003 299 24
ARITHMETICAL ABSTRACTIONS--THE MOVEMENT TOWARD CONCEPTUAL MATURITY UNDER DIFFERING SYSTEMS OF INSTRUCTION.
BY- BROWNELL, WILLIAM A.
CALIFORNIA UNIV., BERKELEY
REPORT NUMBER CRP-1676 PUB DATE JUL 64
REPORT NUMBER OEC-2-10-103
EDRS PRICE MF-$0.54 HC-$14.64 366P.

DESCRIPTORS- *MATHEMATICS INSTRUCTION, *ARITHMETIC, *CONCEPT FORMATION, *INSTRUCTIONAL PROGRAMS, ELEMENTARY EDUCATION, COMPARATIVE ANALYSIS, INTERVIEWS, *MATHEMATICAL CONCEPTS, *COMPARATIVE EDUCATION, *CONCEPT FORMATION, BERKELEY, CALIFORNIA, ENGLAND, SCOTLAND

SEVERAL TEACHING PROGRAMS OF MATHEMATICAL CONCEPTS WERE INVESTIGATED IN AN EFFORT TO DEVELOP MEANS FOR ENHANCING CONCEPTUAL MATURITY AMONG CHILDREN. INTERVIEWS WERE ASSIGNED TO ENGLISH AND SCOTTISH SCHOOLS WHERE DIFFERENT ARITHMETIC PROGRAMS WERE TAUGHT. DATA WERE COLLECTED ON SCHOOL ACHIEVEMENT AND SCHOLASTIC ABILITY FROM INTERVIEWS WITH INDIVIDUAL SUBJECTS. A SERIES OF PRINTED ARITHMETICAL TASKS WAS EMPLOYED TO ASCERTAIN SPECIFIC MENTAL PROCESSES. THE RELATIVE EFFECTIVENESS OF ONE PROGRAM OVER THE OTHER WAS NOT FOUND. FURTHER RESEARCH ON CERTAIN GENERAL PROBLEMS IN ARITHMETIC EVALUATION WAS RECOMMENDED. (RS)

U. S. DEPARTMENT OF HEALTH, EDUCATION AND WELFARE
Office of Education
This document has been reproduced exactly as received from the person or organization originating it. Points of view or opinions stated do not necessarily represent official Office of Education position or policy.

ARITHMETICAL ABSTRACTIONS: THE MOVEMENT TOWARD CONCEPTUAL MATURITY UNDER DIFFERING SYSTEMS OF INSTRUCTION

Cooperative Research Project No. OE 2-10-103

by

William A. Brownell
University of California
Berkeley, California

July, 1964

The research reported herein was supported by the Cooperative Research Program of the Office of Education, U. S. Department of Health, Education, and Welfare

ED003299

APOLOGIA

A personal note to the reader

Because of the analytic nature of this investigation, the report is necessarily long and detailed; and, I fear, overly repetitious.

It is long because a great deal of space is required to present at all adequately the huge body of data collected from more than 1400 children, as well as to present and interpret the findings derived from those data.

The report is detailed because the crucial data were procured individually from children by means of interviews designed to ascertain their thought procedures in dealing with several kinds of arithmetical item, because thought procedures are almost limitless in their variety, and because their integrity had to be preserved as fully as possible.

And the report, unfortunately, is repetitious, for a number of reasons. For one, a writer, like me, can vary his language only so much in expressing essentially the same idea. For another, I have had to provide verbal signposts, in the form of paragraphs reiterating something already said, in order to prevent you, the reader, from becoming confused and losing perspective in the long successions of analyses and comparisons.

The reader who is interested in quantitative data per se and in the way in which in the present instance they were treated to arrive at the conclusions drawn, faces a formidable task in studying the report in its entirety, for there are many tables and many analyses, and -- once again -- the report is long.

Being in full complete sympathy with the reader who is less concerned about such matters and wishing to lighten his burden, I have inserted many summaries throughout the text which are readily identifiable in the Table of Contents. These summaries the more critical and studious reader may pass over, thus avoiding some of the repetitiousness of the report; but they will enable the reader who wants to know the findings only, to get ahead rapidly.

Whatever kind of reader you are, I wish you well.

William A. Brownell

William A. Brownell

ACKNOWLEDGMENTS

It is impossible for me to do what I should like to do and to thank individually all who assisted in making this study. The sheer number of such persons defeats me, even if I exclude the hundreds of children who presented themselves for interviews and the dozens of school officials and of teachers who cheerfully put up with the consequent inconveniences and with the attendant interruptions of their daily schedules.

To all those not named below I can say only that I am very grateful, and that never before have I enjoyed so much cooperation in a research venture. Not a single Director of Education in England and Scotland whom I approached, and not a single headmaster or headmistress of an English or Scottish school declined my invitation to participate in the study. Very many of them displayed genuine interest in the project while it was underway; and not a few, then and since the completion of the study, have expressed their appreciation for the opportunity they had had, in a new way, to learn more about the processes of teaching and learning arithmetic. As I have written elsewhere about this investigation, "One could scarcely ask for more."

I am indebted to the U. S. Office of Education for the grant for the study and to the administration of the University of California, Berkeley, for arranging a leave of absence for me to conduct the inquiry.

The National Foundation for Educational Research in England and Wales and the Scottish Council for Research in Education were exceedingly helpful, first, by lending their prestige to the research by letting it be known that it had their support; second, by offering many valuable criticisms and suggestions; third, by identifying schools that might possibly be used; and, fourth, by making available to me their office facilities and their staff members in emergencies (of which, unfortunately, there were many). I should like to express my sincere gratitude to Dr. D. A. Walker, Director of the Scottish Council, and to Dr. D. W. Wall, Director of the National Foundation, - also, and especially, to Mr. John Biggs of the latter organization, now Dr. John Biggs of the faculty in psychology at the University of New England, N. S. W., Australia. Mr. Biggs was never too busy to respond to my unanticipated calls for his assistance.

Mr. C. R. E. Gillett, of Henley-on-Thames, a former English headmaster, acted as coordinator for the investigation in both English and Scottish schools. I shall always be thankfully indebted to him for his prompt, responsible, and intelligent services, as well as for the cheerful, pleasant manner in which he performed them.

To the authors of the descriptions of the three arithmetic programs studied (Chapter 2), I obviously owe a great deal. Their names are: Miss Margaret L. F. Law of the Moray House College of Education, Edinburgh; Mr. J. Clitheroe, Deputy Director of Education, Carlisle, England; and Mr. L. G. W. Sealey, Adviser to Junior Schools, Leicester, England. Mr. Clitheroe also assisted greatly in setting up in his schools the first of the interview programs; and Mr. Sealey, by suggesting schools where the Dienes system of instruction had been taught for some time and by interesting four of his headmasters in acting as interviewers.

Miss A. Elizabeth Adams, General Inspector of Education, Surrey County Council, put me in her debt by instituting a vigorous search for qualified interviewers and by locating Dienes schools in the area adjacent to London.

Mrs. M. McIntosh, under most adverse circumstances, acted ably as "assistant coordinator" in arranging for interviews in the schools in and near Glasgow, and merits much more than this mere acknowledgment of her help.

My colleague, Dr. Arden Ruddell, of the Department of Education at the University of California, Berkeley, not only participated in the preliminary planning of the investigation but, by working with local children, aided in developing the first drafts of the interview schedules and in re-shaping the interview procedure itself. Besides all this, he read several sections of the manuscript for this report and offered insightful comments and suggestions.

I am grateful to the following residents of Berkeley: Miss Katherine Stauffer, Administrative Assistant in the Department of Education, who, among other things, prevented me from disregarding or fracturing the many regulations emanating from various sources; Miss Elaine Lysne, who worked long hours and performed admirably the functions of research assistant and who carefully scrutinized the typescript for errors; and Mrs. Angelo M. Polese, who patiently and efficiently typed and retyped chapters of the report as they were written and rewritten and saw the manuscript through to its final form.

I come last to the interviewers themselves who furnished me all the data -- and trustworthy data -- used in preparing this report. I am indebted to them beyond measure. Starting with some trepidation, as well they might have, they developed quickly into skillful and intelligent questioners and observers of the children who came to them. It is a pleasure to acknowledge my obligation to them individually as I list their names.

English Interviewers

Mrs. M. L. Baxter,
South Norwood
D. A. Botterill,
Leicester
Miss G. F. Boone,
Wimbledon
Mrs. O. Cleaver,
Luton
Mrs. John Cumber,
Reading
Miss W. Edgar,
Carlisle
Mrs. K. E. Grimwood,
Morden
Mrs. Dorothy A. Jacobs,
Woodley

W. E. Looker,
Leicester
A. E. M. Morris,
Leicester
Mrs. Jane Savage,
A.E.R.E, Harwell
Miss Maureen Shirley,
Carlisle
Mrs. K. Skull,
Streatham
Miss Maude Spedding,
Wandsworth Common
Miss V. Stewart,
Carlisle
Mrs. N. Tait,
Carlisle

Mrs. Pauline Jenkins,
Bovingdon

Miss T. G. Waight,
Reading
H. E. Ward,
Leicester
Mrs. E. L. Whitworth,
Hitchen

Scottish Interviewers

Mrs. Margaret J. Adams,
Kirkaldy
Miss Jean Bain,
Glasgow
Mrs. Janet Campbell,
Edinburgh
Miss J. Dingwall,
Edinburgh
Miss Jean Green,
Glasgow

Mrs. Elizabeht A. McNamara,
Kirkaldy
Mrs. Jean McHugh,
Glasgow
Miss Ivey E. L. Phillips,
Edinburgh
Mrs. Elsie H. Stewart,
Edinburgh

William A. Brownell

William A. Brownell

TABLE OF CONTENTS

Page

Apologia i

Acknowledgments ii

List of Tables ix

List of Figures xiv

Part I. Problems and Procedure 1

Chapter 1. Problems and General Plan of Attack 2

The Problem 2

Abstractions and abstractness 2
Progress toward conceptual maturity 4
Instructional programs and conceptual learning 5
The problem 6
Source of data 7
Arithmetic content 7
Adequacy of samples 8
Evaluation vs fact-finding

General Plan of Attack 11

Subjects 11
Specific arithmetical tasks 11
Types of data 13

Chapter 2. Arithmetical Programs Studied 17

The Cuisinaire (Cui.) Program 17
The Dienes (Dien.) Program 23
The Conventional (Con.) Program 28
Comparisons and Warnings 34

Chapter 3. Details of the Procedures 39

The Training of Interviewers 39
Data and Method of Collection 41
Data on pupils 41
Interview data 42
Classification for Conceptual Maturity 45
Computation 46
Problems 48
Number combinations 50
Treatment of the Data 53

Page

Part II. The Investigation in Scottish Schools 56

Chapter 4. Gross Comparison of the Effectiveness of the Cui. and the Con. Programs 57

Data on Schools and Subjects 57

The Number Combinations 59

Conceptual maturity 59
Proficiency 67
Summary: The number combinations 69

Computation 70

Conceptual maturity 70
Proficiency 73
Summary: Computation 75

Problem Solving 76

Conceptual maturity 76
Proficiency 81
Summary: Problem solving 83

Chapter Summary 83

Conceptual maturity 83
Proficiency 88

Chapter 5. Factors Possibly Influencing the Effectiveness and the Relative Effectiveness of the Programs 90

1. Amount of Instructional Time as a Factor 93

Intra-program comparisons 93
Cross-program comparisons 103
Summary, Section 1 111

2. Brightness as a Factor 114

Intra-program comparisons 115
Cross-program comparisons 123
Summary, Section 2 131

3. General Arithmetic Achievement as a Factor 133

Intra-program comparisons 134
Cross-program comparisons 141
Summary, Section 3 147

Page

4. Reading Achievement as a Factor 151
  Intra-program comparisons 152
  Cross-program comparisons 158
  Summary, Section 4 162

5. Age as a Factor 165

Chapter Summary 166

Part III. The Investigation in English Schools 179

Chapter 6. Cross Comparisons of the Effectiveness of the Cui., the Con., and the Dien. Programs 180

Data on Schools and Subjects 180

The Number Combinations 184
  Conceptual maturity 184
  Proficiency 189
  Summary: The number combinations 192

Computation 192
  Conceptual maturity 192
  Proficiency 195
  Summary: Computation 197

Problem Solving 197
  Conceptual maturity 197
  Proficiency 201
  Summary: Problem solving 202

Chapter Summary 203

Chapter 7. Factors Possibly Influencing the Effectiveness and the Relative Effectiveness of the Programs 207

1. Amount of Instruction as a Factor 209
  Intra-program comparisons 211
  Cross-program comparisons 223
  Summary, Section 1 233

2. Brightness as a Factor 240
  Intra-program comparisons 240
  Cross-program comparisons 250
  Summary, Section 2 260

Page

3. General Arithmetic Achievement as a Factor 266

Intra-program comparisons 267
Cross-program comparisons 277
Summary, Section 3 284

Chapter Summary 290

Part IV. Final Chapter 299

Chapter 8. Findings, Interpretations, Implications, and Related Research 300

Findings 300

The relative effectiveness of the programs 301
Factors possibly influencing the results 303
Conceptual maturity vs proficiency 307

Interpretations 308

The relative effectiveness of the Cui. and the Con. programs 308
Apparent weaknesses in the programs 312
Conceptual maturity vs proficiency once again 315
Objectives, emphasis, pace of instruction 318

Implications 321

Improving the arithmetic program in the first three grades 322
Evaluative research 333

Related Research 339

Appendix 342

The Interview Problems 343

English Study: Tables for Reading Achievement as a Factor 344

LIST OF TABLES

Note: The actual titles of tables are shortened somewhat in the listing below.

Part II. The Study in Scottish Schools

Chapter 4. Gross Comparisons -- the Cui. and the Con. Programs

Table Page

1 General Information Concerning Cooperating Schools 58

2 Comparison of Characteristics of Cui. and Con. Subjects 60

3 Characteristics of Subjects, Weighted Means 61

4 The Number Combinations. Thought Processes with the Eight Combinations 62

5 The Number Combinations. Thought Processes, Operation by Operation 66

6 The Number Combinations. Proficiency: Accuracy 68

7 The Number Combinations. Proficiency: Rate 68

8 Computation. Quality of Explanations 71

9 Computation. Proficiency: Accuracy and Quality of Performance 74

10 Problem Solving. Reasons Given for Solutions 77

11 Problem Solving. Proficiency: Accuracy and Performance Level 82

Chapter 5. Factors Possibly Influencing the Effectiveness and the Relative Effectiveness of the Programs

1. Amount of Instruction

12 Time as a Factor. Composition of the H and the L Groups in Each Program 94

13 The Number Combinations. Thought Processes 95

Table		Page
14	Computation. Quality of Explanations	96
15	Problem Solving. Reasons Given for Solutions	97
16	The Number Combinations. Proficiency	98
17	Computation. Proficiency	99
18	Problem Solving. Proficiency	100
	2. Brightness	
19	Brightness as a Factor. Composition of the H and L Groups in Each Program	115
20	The Number Combinations. Thought Processes	116
21	Computation. Quality of Explanations	117
22	Problem Solving. Reasons Given for Solutions	118
23	The Number Combinations. Proficiency	118
24	Computation. Proficiency	119
25	Problem Solving. Proficiency	119
	3. General Arithmetic Achievement	
26	General Arithmetic Achievement as a Factor. Composition of the H and L Groups in Each Program	134
27	The Number Combinations. Thought Processes	135
28	Computation. Quality of Explanations	136
29	Problem Solving. Reasons Given for Solutions	136
30	The Number Combinations. Proficiency	137
31	Computation. Proficiency	138
32	Problem Solving. Proficiency	138

Table Page

4. Reading Achievement

33 Reading Achievement as a Factor. Composition of the H and L Groups in Each Program 151

34 The Number Combinations. Thought Processes 153

Computation. Quality of Explanations 154

36 Problem Solving. Reasons Given for Solutions 154

37 The Number Combinations. Proficiency 155

38 Computation. Proficiency 155

39 Problem Solving. Proficiency 156

Part III. The Study in English Schools

Chapter 6. Gross Comparisons -- the Cui., the Dien., and the Con. Programs

101 General Information Concerning Cooperating Schools 182

102 Comparison of Characteristics of Cui., Dien., and Con. Subjects 183

103 Characteristics of Subjects, Weighted Means 184

104 The Number Combinations. Thought Processes with the Eight Combinations 185

105 The Number Combinations. Thought Processes, Operation by Operation 187

106 The Number Combinations. Proficiency: Accuracy 190

107 The Number Combinations. Proficiency: Rate 191

108 Computation. Quality of Explanations 193

109 Computation. Proficiency: Accuracy and Quality of Performance 196

110 Problem Solving. Reasons Given for Solutions 198

111 Problem Solving. Proficiency: Accuracy and Performance Level 200

Table Page

Chapter 7. Factors Possibly Influencing the Effectiveness and the Relative Effectiveness of the Programs

1. Amount of Instruction

112 Time as a Factor. Composition of the H and the L Groups in Each Program 210
113 The Number Combinations. Thought Processes 212
114 Computation. Quality of Explanations 213
115 Problem Solving. Reasons Given for Solutions 214
116 The Number Combinations. Proficiency 215
117 Computation. Proficiency 216
118 Problem Solving. Proficiency 216
118A Time as a Factor. Relative Standing of Groups in Conceptual Maturity 234
118B Time as a Factor. Relative Standing of Groups in Proficiency 239

2. Brightness

119 Brightness as a Factor. Composition of the H and L Groups in Each Program 241
120 The Number Combinations. Thought Processes 242
121 Computation. Quality of Explanations 244
122 Problem Solving. Reasons Given for Solutions 244
123 The Number Combinations. Proficiency 245
124 Computation. Proficiency 245
125 Problem Solving. Proficiency 246
125A Brightness as a Factor. Relative Standing of Groups in Conceptual Maturity 261
125B Brightness as a Factor. Relative Standing of Groups in Proficiency 265

Table Page

3. General Arithmetic Achievement

126 General Arithmetic Achievement as a Factor. Composition of the H and L Groups in Each Program 267

127 The Number Combinations. Thought Processes 268

128 Computation. Quality of Explanation 269

129 Problem Solving. Reasons Given for Solutions 269

130 The Number Combinations. Proficiency 271

131 Computation. Proficiency 271

132 Problem Solving. Proficiency 272

132A General Arithmetic Achievement as a Factor. Relative Standing of Groups in Conceptual Maturity 285

132B General Arithmetic Achievement as a Factor. Relative Standing of Groups in Proficiency 288

4. Reading Achievement

133 Reading Achievement as a Factor. Composition of the H and L Groups in Each Program 344

134 The Number Combinations. Thought Processes 345

135 Computation. Quality of Explanations 346

136 Problem Solving. Reasons Given for Solutions 346

137 The Number Combinations. Proficiency 347

138 Computation. Proficiency 347

139 Problem Solving. Proficiency 348

Page

LIST OF FIGURES

Fig. 1. The Cuisinaire rods. 20

Fig. 2. Colour families, illustrating the process of progressive doubling. 20

Fig. 3. Use of the Cuisinaire rods to discover the composition of 10. 21

Fig. 4. Two ways of showing 2 x 6 = 12 and 3 x 4 = 12 with the Cuisinaire rods. 21

Fig. 5. Use of the M. A. B. to add 1232 and 233, to get the sum, 2121; base of 4. 26

Fig. 6. Step-by-step development of the domino pattern for 5. 31

Fig. 7. Misrepresentation of the relationship between the Cui., Dien., and Con. programs. 180

Fig. 8 Better representation of the relationship between the three programs. 180

PART I

PROBLEMS AND PROCEDURE

CHAPTER I

PROBLEMS AND GENERAL PLAN OF ATTACK

The Problem

Abstractions and abstractness

Instruction in arithmetic is expected, among other things, to help children acquire a wide range of mathematical concepts and understandings, and a functional knowledge of many mathematical relationships, principles, and laws. These learning products are abstractions, derived from numerous relevant experiences; but there are many degrees of abstractness, so that it is meaningless to say that any child "has" this or that concept or understanding. It is more accurate to say that he "has" the concept or understanding to a greater or less degree of completeness and precision.

An illustration or two will make this point clear. Consider first what Child A and Child B do when they need the answer for 9 + 7.

Child A can find the sum only by manipulating and enumerating objects one by one. Child B, on the other hand, immediately and consistently responds "Sixteen." Moreover, he does so without recourse to enumeration of objects and is able to demonstrate the correctness of his answer by saying, "Ten and six are sixteen; so nine and seven are sixteen;" or "Nine and seven is the same as eight and eight; sixteen;" or "Nine and nine are eighteen; and this is two less; sixteen."

The difference between the levels of maturity and abstractness exhibited by Child A and by Child B are self-evident. Child A cannot think of 9 and of 7 as such, but must reduce each number to a collection of discrete items. Yet, it cannot be said that he is totally without any understanding of the fact 9 + 7 = 16. It is nearer the truth to say that he has attained some degree of abstractness in dealing with it, for he imposes on external objects something they do not themselves possess. The number sequence he repeats partakes of the nature of abstractness. By contrast, Child B, who could have enumerated objects, has abandoned this primitive approach and has gone far beyond Child A in his method of conceiving the numbers involved, and of relating them.

Consider next the ways in which Child C and Child D obtain answers for the problem, "Jim had 10 marbles, but he lost 3 of them. How many marbles were left?" Child C says, "Seven." Asked how he knows, he says, "As soon as I saw (or heard) the word 'left', I knew I should take three from ten." To him, "left" is a subtraction cue, just as are the sign "-" and the word "subtract," encountered over and over again in examples and computational exercises. True, his cue will often mislead him, as, for instance, in attacking the problem, "After eating their candy, Ned had 4 pieces left, and Joe had 2 pieces left. How many pieces of candy were left?" Yet, who is to say that Child C has no conception whatsoever of subtraction as a mathematical operation?

For the problem about marbles above, Child D, who also quickly announces "Seven" as his answer, relies upon no such specific cue as "left" or "lost." He is able to explain his choice of operation by pointing out that, having a total (10) and one of its components (3), he knows that he should subtract to find the other component. Or, he may state that the sense of the problem as a whole requires that the answer should be less than the original total; hence, he should subtract -- an explanation which later on he will need to modify.

If we can say that Child C is at a beginning stage of learning, then we can say that Child D is several steps in advance of him in arriving at a mature grasp of the function of subtraction. Child D's use or non-use of this operation is independent of the presence of such tangible helps as particular words. He has made substantially more progress than has Child C toward mathematical abstractness in this one instance.

Progress toward conceptual maturity

In the foregoing contrasts of Child A and Child B in the case of the number fact 9 + 7 = 16, and of Child C and Child D in the case of problem solving, pairs of children are described who are, as it were, toward the extremes of conceptual learning. Instances of intermediate levels of conceptual learning could be multiplied. For example, the child who finds the answer for 9 + 7 without objects and by silent partial counting ("Nine, ten, eleven, ... sixteen") is further along than is Child A but is still far short of the status exhibited by Child B.

Concepts and understandings are, therefore, not to be thought of in terms of all or none, as being either present or absent. Instead, we must think of the process of acquiring mature and functional concepts in terms of the reorganization of thinking as new experiences afford new insights. Children attain one level of thinking, only to overlay it with a more advanced level, and the latter, with a still more advanced level. Progress in conceptual learning proceeds step-wise; but it can stop at any point.

More than that, the movement toward abstractness and maturity of concepts in the case of any given child is uneven. The mere fact that he has achieved a considerable degree of advancement in one aspect of arithmetic is no assurance that this advancement is general and will be in evidence in other aspects. For example, a child may have mature concepts of the whole numbers 0 to 10 and still know next to nothing about the decimal character of our system of number notation.

Instructional programs and conceptual learning

As has already been stated, progress toward abstractness and maturity of arithmetical concepts (as of all concepts) is dependent upon the kinds and amounts of experience afforded the learner. It follows, therefore, that dissimilar instructional programs, since they provide unlike amounts and kinds of experience, may vary considerably in the degree to which they encourage the acquisition of more and more mature concepts.

Thirty years ago it was common to give children number facts like 9 x 7 = 63 and to practice them interminably in saying and writing the answers. Their experiences were repetitious, and learning consisted in memorizing the corresponding verbal formulas; e.g., "Nine times seven are sixty-three." With the many facts to be memorized, each independently of others, there is little wonder that children forgot or confused answers, and little wonder that the facts, once taken from the context of pure recall, were useless in practical situations. Many a child subjected to such teaching was unable to find the correct answer for a problem like "If you put 7 apples in a bag, how many apples will you need for 9 bags?" Nowhere in the language of the problem could he identify the simple, familiar fact, 9 x 7 = 63.

Of course, what happened was that many of the brighter children, defeated in their attempts to learn the number facts by repeating empty verbalizations, struck out on their own, found relationships among the facts for themselves, and eventually achieved intelligent, functional mastery of the facts. But their progress toward conceptual maturity, be it noted, would never have occurred if they had persisted in verbatim practice, - this and nothing else.

Rather, they had to have different kinds of experience, even if they had to manufacture them for themselves.

Nowadays the arithmetic program provides children the experiences needed for conceptual advancement. The number facts are not given to children; instead, they are found by them. Number relationships are not withheld; instead, experiences are arranged so that they will be "discovered" by learners. Memorization of meaningless formulas is discouraged, not merely because memorization is premature, but because it is detrimental to the development of more mature and useful understandings: it leads children to stop too soon, where they are. While mastery of the number facts is desired and is universally sought as an objective, it is properly viewed as the end result of progressive advances in thinking.

The two instructional programs just described, differing as they do in the learning experiences they afford, should produce quite dissimilar pictures of the progress children make toward maturity of concepts and degree of understanding of the relationships represented in the number facts.

The problem

The purpose of the present investigation was to ascertain what happens to children after being exposed for three years to different instructional programs; that is, what happens in as far as progress toward abstractness and maturity of arithmetical concepts is concerned.

In addition to the major problem, several subsidiary problems will be considered. For one thing, it was noted in the course of the research that schools, even within the same program group, allotted wide variations of time to arithmetic. It is of interest to know the effects, if any, produced by these differences. It is of interest also to know the relationship between rate of conceptualization on the one hand and such factors as comparative age, brightness, and scholastic achievement on the other.

Source of data

Data were collected from a total of 1,406 British children in the spring of 1962. Of this number, 478 were enrolled in Scottish schools, where about half had been taught arithmetic according to one or the other of two programs. The remaining 928 subjects were drawn from English schools, and about a third of them had been taught in accordance with one or another of three arithmetic programs. Details concerning these programs will be set forth in Chapter 2. Until then, it will suffice to designate the programs merely as X, Y, and Z. Programs X and Y were common to Scottish and to English schools, while program Z was to be found only in English schools.

The reasons for conducting the inquiry in the United Kingdom were as follows. Program X, while it had been in use in relatively few schools in the United States prior to 1962, had been in much wider use and for a substantially longer period of time in English and Scottish schools. As a consequence, British teachers who followed this program could, it was assumed, be expected to be more expert in teaching it than could American teachers. Program Y was, and is, practically unknown in this country and in 1962 was available for study only in English schools. Parenthetically, Program Z resembles in some degree programs generally prevalent in the United States for the last two or three decades.

Arithmetic content

Since it was impossible to check conceptual development with respect to all phases of arithmetic, a choice had to be made. Correspondingly, research efforts were confined to attempts to discover

(a) how children taught by each of the three programs arrived at answers for eight number combinations (two in each of the four operations);

(b) how well they understood the mathematical rationale of computation (four examples, one in each operation), and

(c) how they went about the task of solving verbal problems (twelve problems, three in each operation).

Data for the Scottish and the English subjects will be treated separately, thus yielding two parallel but independent investigations, both of them revealing some of the effects of instruction by what have been named Programs X and Y (but not Program Z).

Adequacy of samples

Within each type of arithmetical task the number of items is small. Skill, or lack of skill, in computing answers for the four selected examples hardly affords a basis for judging a given child's general proficiency in computation. He might fail on only these four examples in a test containing 46 other examples, in which case his per cent of accuracy would be 96 instead of 0 if the "test" were limited to just these four examples. In similar fashion, the degree of proficiency shown in dealing with but eight number combinations can provide no assured estimate of proficiency in dealing with the 390 combinations usually taught; and the amount of success achieved by a child in solving the twelve verbal problems in this study tells one little about his probable success in solving dozens of other problems.

None of the three samples is, then, adequate to assessing over-all proficiency in a given type of arithmetical task. However, in this investigation measures of proficiency are of relatively minor significance. Instead, -- and the statement bears repitition -- the chief concern may be stated variously as degree of conceptual maturity, level of thought processes, grasp of mathematical principles and relationships, extent of meaningful arithmetical abstractions. And, in the light of this special concern each sample of items *is* adequate.

It is a reasonable hypothesis that children who can explain the mathematical rationale of their computations for the four examples chosen, one in each operation, will demonstrate about the same degree of ability to explain their computations in other examples they have studied; that children who make use of high (or low) levels of thought processes with the eight rather difficult number combinations put to them in the interviews will use thought processes of the same kind with other equally difficult combinations; and that children who show considerable (or little) ability to explain their solutions of the twelve problems they were asked to solve will be about equally skillful in explaining solutions for other problems of the same kinds.

An additional safeguard with regard to the adequacy of the samples is the fact that the comparisons to be reported are not for this child with that child, or for these ten children with those ten children, but for this large number of children with that large number of children. And these comparisons are to be made in terms of average per cents, not of individual scores. The data on the relative proficiency of the groups compared were collected chiefly in order to ascertain whether advances in conceptual maturity are, or are not, likely to be purchased at a cost to proficiency in the arithmetical tasks employed in the interviews.

Evaluation vs fact-finding

It is not intended now -- and it never was intended -- to evaluate the three programs as wholes, so as to be able in the end to array Programs X, Y, and Z according to their relative merits and to say something like this: Program X is the best, Program Y is the poorest, and Program Z is of intermediate quality and effectiveness, and this order holds for instruction in all phases of arithmetic as taught in the first three years. Such an over-all evaluation can rarely, if ever, be justified, and it probably should not be attempted.

In the first place, no investigation (least of all this one) can possibly lay claims to being so all-inclusive in its scope as to have valid and reliable measures for every phase of arithmetic as it has been taught. If there ever should be such a comprehensive study, the writer wishes the investigator well in the well-nigh impossible task of weighing properly the superiority of one program at this point and its inferiority at other points.

In the second place, the effects of differences in program objectives must be taken into account, a fact which is not at all frequently recognized in evaluative research. Program X, for example, may be designed to produce early mastery of the number facts, with the consequence that it provides the sequence of experiences necessary to a relatively rapid pace of conceptual learning. On the other hand, in Program Y such early mastery of the number facts may be a matter of little importance, progress in learning being left to time, or to chance, or to the concomitant influences of other learning. In these circumstances, in Y-schools a relatively poor showing on the number facts could properly be accepted with complacency, as being of little significance in comparison with progress toward other objectives deemed more important, -- indeed toward objectives which may have been slighted in the research design.

This investigation was planned primarily to collect facts, - to see what happens to children taught according to the various systems of instruction and to see what happens to them chiefly with respect to the movement toward conceptual maturity. Evaluations of the programs will be undertaken almost wholly in terms of their relative effectiveness in this aspect of learning arithmetic. Yet, it was hoped that the facts procured might add to understanding of the process of concept formation and at the same time contribute something to the improvement of instruction in arithmetic in the early school years.

General Plan of Attack

Subjects

As has already been stated, the subjects in the investigation were just completing their third year in school. A larger number of them fell within the age span 7 years 6 months and 7 years 11 months than in any of the other three age categories. As a whole, they were approximately a year younger than typical American third-graders, the reason being that in the United Kingdom children enter school at age 5 instead of age 6. On this account, no valid comparisons can be made between the British subjects and American third-graders. This statement would be true even if data for American third-graders were available, as they are not. We simply do not know how to assess precisely the effects that may be produced by the difference of a year in age.

Specific arithmetical tasks

It was of course essential that the instrument used as the means of procuring data contain only arithmetical tasks representing outcomes expected after three years of schooling in the United Kingdom. The tasks tentatively adopted after preliminary use with American third-graders and after three revisions were submitted to English teachers, school officials, and research workers for their criticisms. Numerous changes were made in order to insure that the tasks would be appropriate for British schools, and the process of revision, based upon try-outs with English children, was continued until this goal seemed to have been achieved.

The instrument as finally arrived at contained three types of arithmetical item. These are:

A. Eight simple number combinations

$$\begin{array}{r}15\\ -\ 7\\ \hline\end{array}\quad \begin{array}{r}4\\ \times\ 6\\ \hline\end{array}\quad \begin{array}{r}8\\ +\ 5\\ \hline\end{array}\quad 7/\overline{21}\quad \begin{array}{r}6\\ -\ 9\\ \hline\end{array}\quad 8/\overline{32}\quad \begin{array}{r}17\\ -\ 8\\ \hline\end{array}\quad \begin{array}{r}9\\ \times\ 3\\ \hline\end{array}$$

B. Four computations

$$\begin{array}{r}28\\ \times\ 3\\ \hline\end{array}\quad \begin{array}{r}92\\ -\ 45\\ \hline\end{array}\quad 4/\overline{84}\quad \begin{array}{r}27\\ +\ 46\\ \hline\end{array}$$

C. Twelve verbal problems, of which only two are printed here. (The twelve problems appear as Exhibit 1 in the Appendix.)

> Once there was a boy called Sam. He had 16 eggs in a basket. 7 of them were cracked. How many eggs were not cracked?
>
> We use 5 sheets of paper to make each of our books. How many sheets of paper will we need to make 5 books?

The number combinations. The two addition combinations and the two subtraction combinations, involving sums of at least 13, are among those with which American children have their greatest difficulty; but their use was justified by the fact that they are generally taught in the first three years in British schools. The two combinations in each of multiplication and division have no products greater than 32 for the reason that in many British schools instruction at the end of the third year stops with the tables involving 4 as a factor. (Actually, as the investigation progresses and as the data to be presented later will show, in some schools instruction was not carried even that far.)

Computation. There was but a single example in each operation. The total sample in computation was therefore very limited; but, as has been explained above, the four examples are adequate for the purposes they are to serve. They involved only two-digit numbers, except in answers, again on the advice of the English authorities consulted; and regrouping or renaming ("carrying" and "borrowing") was limited to the examples in addition, subtraction and multiplication. No division example of the type $3/\overline{135}$ could be included because instruction in the first three years was said to be limited to the type $4/\overline{84}$. As a matter of fact, it was found that in a substantial number of British schools instruction even on this simple type had not been carried to the point of mastery, and in other schools had barely been introduced.

Verbal problems. Some of the twelve verbal problems as first presented to English critics were quite unsatisfactory, partly because of differences in terminology (e.g., "biscuits," not "cookies") and partly because of the relative unfamiliarity of the problem settings. Consequently, a group of five English teachers was asked to prepare appropriate problems, making whatever modifications thought to be necessary in the originals, save with respect to the numbers and the operation to be used in each problem. Here no alterations were allowed.

Types of data

Two distinctly different kinds of data were collected. (a) The one type afforded evidence on proficiency: correctness of answers, quickness of response; and (sometimes) general quality of performance, such as steady, smooth performance vs erratic or uneven performance. These measures were obtained partly from visual observation and partly from listening to what each child said.

(b) The second kind of data was intended to describe and classify (1) what each child thought as he dealt with each number task or (2) how much he understood what he had done. Intimate and valid information of this sort is not to be had from written work or even from what the child says voluntarily, apart from questioning. Inferences are fraught with the danger of error. On this account, the technique employed was that of interviews with individual children.

Permanent records were made possible by the use of large printed schedules, one for each child, which the interviewer had before him as he observed and questioned children. As the result of considerable prior interviewing, both in American and in English schools, a code of symbols was devised, so that, for most responses, interviewers needed only to encircle or check. Space was left, however, for the recording of interesting responses and of responses interviewers believed they could not classify accurately. These written statements proved invaluable by furnishing new information and by enabling the writer to assign ambiguous responses to their correct categories.

With the exception of the computational exercises, children were not called upon to write answers or to read statements concerning arithmetical tasks. The verbal problems and the number combinations were all presented orally. More complete information concerning the conduct of the interviews will appear in Chapter 3.

The interviewers

Despite its seeming simplicity, interviewing is actually most complicated and difficult. It is especially so when its purpose is to secure accurate and reliable information of the kind needed in this investigation. Thought processes are evanescent, and may not be exactly recalled; understandings may not always be readily verbalized; and children may be suggestible or uncooperative under questioning. Hence it was imperative that only properly qualified interviewers be employed.

The original plan was to organize two cadres of interviewers, about six or seven for Scottish schools and eight or nine for English schools, to train them carefully and thoroughly in Edinburgh and in London, respectively, and then to have them move about from school authority (district) to school authority. This arrangement proved to be impracticable, for too few persons could be found who were of the kind desired and who were free to do the required travelling. The arrangement would have been unworkable in any case. As was quickly learned, such peripatetic interviewers might have been somewhat less than welcome in authorities with which they had had no prior relationship. It became necessary, therefore, to find and train separate corps of interviewers for each school authority (or, in a few instances, for a group of adjacent school authorities). This change in plans meant that more interviewers than had been anticipated had to be recruited and that much more time had to be given to preparing them for their work.

But for the extraordinary cooperativeness and helpfulness of local teachers and school officials, the whole investigation could have collapsed in the face of these problems. To these persons in each cooperating authority the qualifications of interviewers were explained. (a) They should have had at least several years of successful experience in teaching in the Infant School and/or in the Junior School. (b) They should be patient and skilled in working with children in a relaxed, informal, and friendly atmosphere. (c) They should be imaginative, so as to ask the right questions at the right time, and flexible, so as to be able to change directions as needed in the questioning. (d) They should be able to learn quickly what they were expected to do. (e) They should have good judgment in order to discriminate correctly among types of categorized responses. And (f) they should be known to be habitually careful and thorough, as a guarantee that they would record correctly and completely all entries on interview schedules.

In every school authority enough such interviewers were quickly found, within a day or two except in one instance. In all, they numbered 29. Of the nine Scottish interviewers five were headmistresses at the time, and one was a former headmistress. Of the twenty English interviewers seven were headmasters or headmistresses, one was a deputy headmistress, and one was a former headmistress. The other fourteen interviewers, who had had no administrative experience, nevertheless fully satisfied the qualifications listed above. The fact that thirteen headmasters and headmistresses (including the deputy headmistress) chose to participate in the study was as gratifying as it was unexpected. Having learned about the purpose of the research and about the nature of the procedures to be employed, they were eager to serve as interviewers. As one headmistress put it later, "For years I had been hoping for a chance like this, to talk with children in an unhurried way and to find out how they think. I wouldn't take anything for this experience."

CHAPTER 2

ARITHMETICAL PROGRAMS STUDIED

In order to assure authenticity, statements concerning the three arithmetical programs, heretofore designated merely as Program X, Y, and Z, were written by three persons, each known to be exceptionally qualified for his assignment. These persons, all of them citizens of the United Kingdom, were asked, in limited space, to tell about the program he knew best by reason of study and/or first-hand experience. Their statements follow, together with additional information elicited through correspondence. The gratitude of the writer to these "conscripted" cooperators is very great indeed.

The three programs will be referred to as the Cuisinaire (Cui.) program, the Dienes (Dien.) program, and the Traditional or Conventional (Con.) program.

The Cuisinaire (Cui.) Program

The statement regarding the Cuisinaire (Cui.) program was prepared by Miss Margaret L. F. Law, M.A., Edinburgh, and Senior Lecturer in Educational Methods at Moray House College of Education, Edinburgh. Among her other qualifications, Miss Law studied with the founder of the Cuisinaire system of instruction, Monsieur Georges Cuisinaire, in his schools in Belgium and is frequently sought as a lecturer or consultant in schools which have adopted the Cui. program or are anticipating doing so. She has just published materials for pupils and teachers, to provide "a programme which should take children through the first three years of their mathematical studies, based upon a fairly wide approach which would be acceptable to the majority of teachers."

M. Cuisinaire, described by Miss Law as "a primary teacher who dealt with all types of children and was well aware of the differences in level of intelligence with which the class teacher has to cope," and as "the least dogmatic of men" (a judgment corroborated by the writer in conversations with many British school people who had talked with him or corresponded with him), -- M. Cuisinaire, who had built on the work of Tillich and others, gave "the world contract of [his] material" to Dr. Caleb Gattegno. The latter, a "continental mathematician," to use Miss Law's words, "who up to the time had worked mostly with university students or children with at least average ability and not much with really young children," saw "the mathematical possibilities in the material and greatly enlarged its scope, being struck with the opportunities it presented for generalization." Accordingly, Dr. Gattegno stressed algebra more, and arithmetic less, than had M. Cuisinaire; and he formulated a system of instruction to which British teachers who follow the "Cui. program" adhere more or less scrupulously.*

Miss Law's description of the main features of the Cui. program follows in her own words.

"The Cuisinaire material consists of a box of 241 coloured wooden rods each having a cross-section of 1 sq. cm. These rods increase in length by 1 cm. at a time, from the smallest unit which is a 1 cm. cube to the longest rod which is a prism 10 cm. long. (Fig. 1)

---

* Miss Law has listed what she regards as the most useful references on the Cuisinaire program, or on the Cuisinaire-Gattegno program. The first four have been published since 1959. The fifth is in press at this writing.

G. Cuisinaire and C. Gattegno. Numbers in Color. London: William Heinemann, Ltd.

C. Gattegno. Arithmetic with Numbers in Color. Books 1 to 10. Reading, England: The Cuisinaire Co.

C. Gattegno. A Teacher's Introduction to the Cuisinaire-Gattegno Method of Teaching Arithmetic. Reading, England: The Gattegno-Pollock Education Co., Ltd. Now to be had from The Cuisinaire Co. as above.

C. Gattegno. Now Johnny Can Do Arithmetic. Reading, England: Educational Explorers, Ltd.

Margaret Law. Number Building. Teacher's Manual, with Texts. London: Thomas Nelson and Son. Ltd.

"The red (2), crimson (4), and brown (8) rods form the first colour family with 2 as the common factor. The yellow (5) and orange (10) rods form the second family with 5 as the common factor. The light green (3), dark green (6), and blue (9) rods form the third family with 3 as the common factor. It will be noticed that the colour deepens within each family as the numerical value of the rods increases. The white and black rods represent the prime numbers 1 and 7 respectively.

"There is also a wall chart depicting in colour families the products of numbers within the hundred when neither of the two factors exceeds ten. These products are obtained through a process of progressive doubling starting from the smallest family unit. (Fig. 2)

"Lotto cards are provided, each bearing a selection of the product pictures, together with matching single-product cards and numbered counters. These counters are placed in the blank centre circle of the appropriate product picture when games are played. The chart and the cards are a pleasant means of impressing the multiplication facts (discovered through the use of the rods) on the child's memory.

"Operating with the rods the child is presented with concrete pictures whereby he can compare quickly and accurately different numbers. By placing below any given rod smaller rods which, end to end, make up its length, he can discover the compositions of its value. (Fig. 3)

"The situation where the length of the rod is composed with rods of one colour only demonstrates the principles of multiplication and division.

"The introduction of fractions presents little difficulty, and this aspect of number relations can be tackled at a much earlier stage than was formerly possible. Looking at the multiple picture the child reads it as 2 x 5 and sees the whole as made up of two equal parts. In his everyday experience he knows that each of these parts can be called 'one half' and he can therefore read from the same picture $5 = \frac{1}{2}$ of 10.

"Block pictures of the products made by placing equal rods side by side instead of end to end can be symbolically portrayed by crossing one of these equal rods by a rod representing the number of times we have used it in making our rectangle, thus providing an economical form of rod notation. (Fig. 4)

"These block pictures lead easily to the study of squares and cubes, while the building of towers of crossed rods prepares the way for the study of indices. An introduction to the study of algebra can be achieved through using the initial letters of the colours to form equations.

"It is the usefulness of this material at all stages of the mathematical programme that has ensured its acceptance by so many practising teachers, to whom its qualities of simplicity and flexibility are immediately apparent. It gives opportunity for continuing purposeful activity on the part of the child, as a result of which not only are the number bonds easily perceived and memorized, but an understanding of mathematical relationships is engendered."

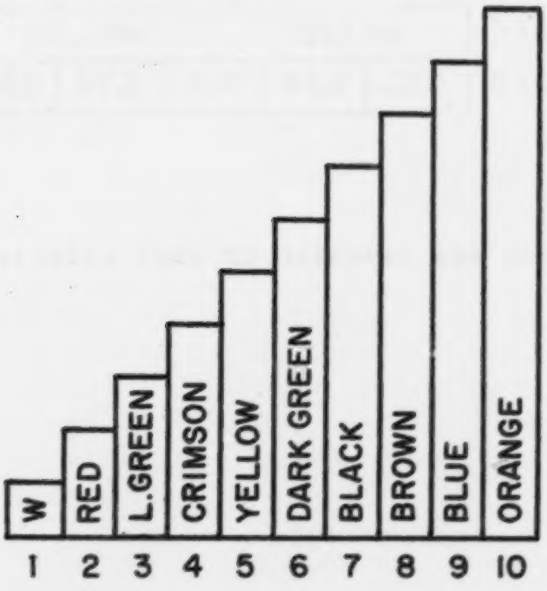
W
RED
L.GREEN
CRIMSON
YELLOW
DARK GREEN
BLACK
BROWN
BLUE
ORANGE
1
2
3
4
5
6
7
8
9
10

Fig. 1. The Cuisinaire rods.

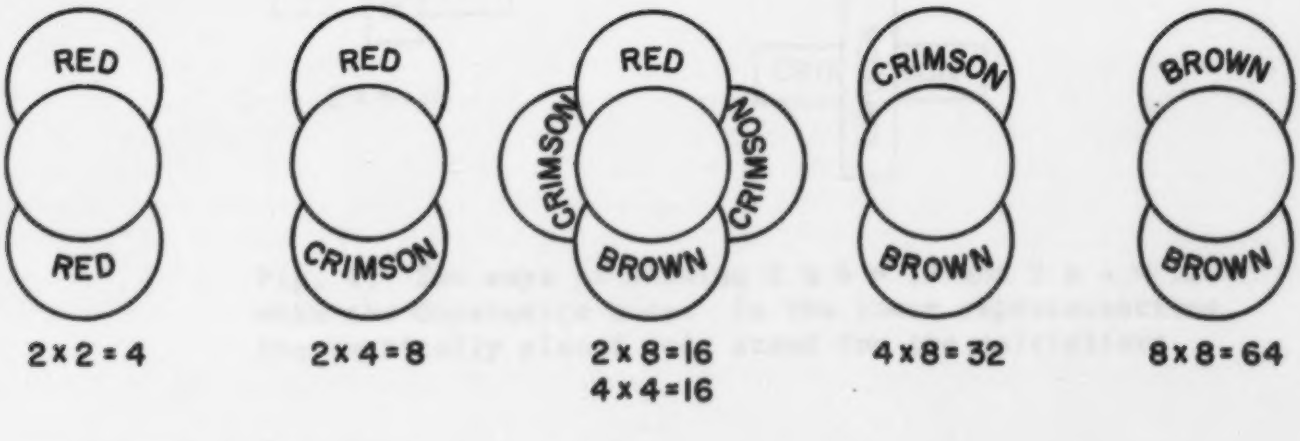
RED
RED
2×2=4
RED
CRIMSON
2×4=8
RED
CRIMSON
CRIMSON
BROWN
2×8=16
4×4=16
CRIMSON
BROWN
4×8=32
BROWN
BROWN
8×8=64

Fig. 2. Colour families, illustrating the process of progressive doubling.

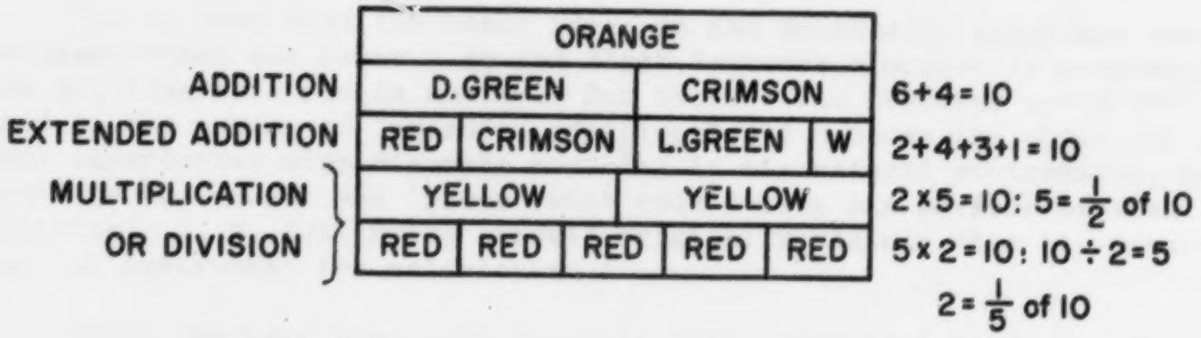
ORANGE
ADDITION
D. GREEN
CRIMSON
6+4=10
EXTENDED ADDITION
RED
CRIMSON
L.GREEN
W
2+4+3+1=10
MULTIPLICATION
OR DIVISION
YELLOW
YELLOW
2×5=10: 5=$\frac{1}{2}$ of 10
RED
RED
RED
RED
RED
5×2=10: 10÷2=5
2=$\frac{1}{5}$ of 10

Fig. 3. Use of the Cuisinaire rods to discover the composition of 10.

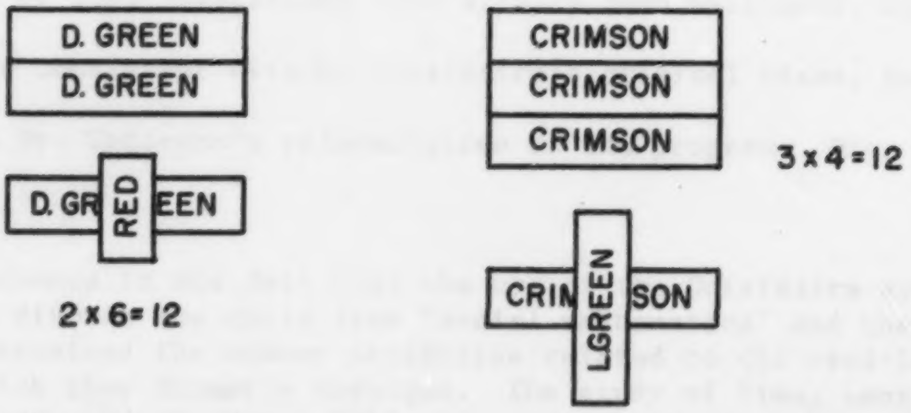
D. GREEN
D. GREEN
D. GREEN
RED
2×6=12
CRIMSON
CRIMSON
CRIMSON
3×4=12
CRIMSON
L.GREEN

Fig. 4. Two ways of showing 2 x 6 = 12 and 3 x 4 = 12 with the Cuisinaire rods. In the lower representations the vertically placed rods stand for the multipliers.

Additional information

In order to clarify certain issues, three questions were put to Miss Law by corraspondence. The questions and her answers, necessarily abbreviated, are as follows:

(a) Is it fair to say that the Cui. program emphasizes ratio and measurement, rather than experiences with counting and then grouping discrete objects?

"It is true that the basic ideas in the Cuisinaire programme are those of measurement and ratio. In the first instance a number is presented to the children as a single entity. But because the discrete units will be used in the making of 'extended additions' and because the child has had many experiences with discrete counting in his natural environment, he may continue to 'count' and 'group' with rods during his periods of free activity. ... M. Cuisinaire himself originally claimed that his material was 'an instrument for calculation.'

"Many teachers have felt strongly that 'counting' is a natural approach to mathematics and that its importance cannot be denigrated ... Other types of apparatus employed ... are counters, pictorial group representations of the quantitative value of the numbers 1 - 10, and material designed to augment the child's experience in space perception .... [Some teachers] make identification of the quantitative value of any number [a necessary step] in the discrete situation, using materials the child handles in the process of living," e.g., number of seats in a bus, number of passengers to sit in them, "before introducing the Cuisinaire material to children for the formal study of mathematics."

(b) To what extent, if at all, and how, are activities other than those with the Cuisinaire rods stressed in Cui. schools?

Some types of supplementations have already been mentioned, types which, be it noted, are consistent with M. Cuisinaire's original ideas, but not too consistent with Dr. Gattegno's reformulation of the program. Miss Law goes on to say,

"In many schools it was felt that the use of the Cuisinaire apparatus was inclined to divorce the child from 'social mathematics' and the teachers have therefore retained the number activities related to the real-life situation in which they formerly indulged. The study of Time, Length, and Weight Measurement, and the Money Tables is introduced throughout the curriculum as the need to know more about such topics arises in the pursuance of the children's interests, and is mostly attacked in the first instance in the practical situation; e.g., shopping, measuring with scales."

Miss Law estimates that "in Scotland and in many English schools 15 - 20% of the regular class period is devoted to such activities;" that is to say, activities which do not involve use of the Cuisinaire rods.

(c) Is it fair to say that in Cuisinaire schools early mastery of the simple number facts is not regarded as a major objective?

"I think a fairer statement would be that the Cuisinaire programme is not solely concerned with the mastery of the number bonds, or facts, or combinations. There is evident among a considerable amount of teachers a marked feeling that these are still useful and important and that their memorization is not incompatible with the development of insight into mathematical relations. There are also those who feel that it is to the recognition of 'mathematical relationship' that the child will come naturally and in his own good time." ...

M. Cuisinaire has suggested "a procedure to be used in directed lessons on the study of any number, which he feels will assist the child towards the memorization of the basic number facts." By contrast, Dr. Gattegno, whose variation of the Cui. program seems to be in common usage, "feels that 'memorization' will come through 'operation'," that is to say, through use.

The Dienes (Dien.) Program

The term used to designate the second of the arithmetic programs investigated is not altogether a happy choice. It has been named for the man who brilliantly developed a theoretical framework for the teaching of mathematics, including arithmetic, and who devised a most elaborate and a most ingenious body of apparatus to be used in connection with his views. In the late 50's Dr. Z. P. Dienes, then Lecturer in Mathematics at the University of Leicester (he is now reader in Psychology at Adelaide University, South Australia), began to work with children in nearby schools, to test his theories and the usefulness of his apparatus. From the outset, however, his attention was centered chiefly on the learning and the teaching of mathematical laws, generalizations, and the like; and his interest in arithmetic, though real, was nevertheless secondary to his interest in mathematics. Moreover, while some of his apparatus was found to be suited to children in the second year of the Infant School, this apparatus and other forms of his apparatus proved to be of greater value later on in the Junior School. Hence, the Dien. subjects in this research had had but limited experience with Dr. Dienes' ideas and concrete aids save in their third year of schooling.

From the beginning, Mr. L. G. W. Sealey, then and now Advisor to Junior Schools in Leicestershire, was Dr. Dienes' principal collaborator. But, unlike Dr. Dienes, he had to be concerned that the children in the experimental schools should be arithmetically competent; and he took steps to guarantee that this should be so. The name "Dienes Program" is therefore misleading, for it tends to minimize the extent of the arithmetical content which, due to Mr. Sealey's efforts snd leadership, made up and still makes up the larger part of the program for the first three years in Dien. schools. Nevertheless, the program will continue to be called the Dienes program, for there is no other name which can give proper recognition to the source of the really unique features of instruction beginning usually at the middle or near the end of the second Infant School year.

Mr. Sealey has prepared the description of the Dien. program. In his statement he stresses the particular value of Dr. Dienes' Multibase Arithmetical Blocks (the M. A. B.); but he then points our their limitations, depicts the need for arithmetical experiences of other kinds, and explains what is done in Dien. schools to provide them.

Mr. Sealey is widely known throughout England at present as the outstanding authority on the Dien. program (as that term is used in this report). He is in great demand as a lecturer and as the director of workshops where the program has been adopted or is being considered for adoption. He has also been a prime mover in setting up the program in many schools, several of which in the area south of London participated in this inquiry.

For reasons that can now be understood, Mr. Sealey entitles his statement, "The Multibase Arithmetical Blocks (M. A. B.) in Leicestershire Schools." In order to fit it into the present context, a few slight, but easily detectable alterations have been made to his text, and there has been added information gained through correspondence with Mr. Sealey. In view of these changes he was asked to read and approve the material here subsumed under the name, the Dien. program, and this he has been good enough to do.

"The M. A. B. material consists of wooden items commonly known as 'units,' 'longs,' 'flats,' and 'blocks.' The pieces represent the powers of various number scales, the 'long' representing the base, or first power. The scales chosen by Dr. Dienes were those to the base of 3, 4, 5, 6, and 10. The unit, or zero power, is common. It is actually a three-eighths inch edge cube. Thus, in the 4-base set, the 'long' is a marked wooden rod equivalent to 4 of the units, the 'flat' is equivalent to 16 units, and the 'block' is equivalent to 64 units. [See Fig. 5 and the accompanying explanation.]

"The material was first made in 1958 and introduced into two primary schools in Leicestershire at this time. Since then, almost half of the primary schools in the County have obtained the material and use it regularly as an essential part of the young child's mathematical learning experience. In no case does work with the M. A. B. constitute the total mathematical experience of any child. Indeed, this would not be possible.

"The M. A. B. is so structured that it leads a child to an understanding of the processes of addition and subtraction and multiplication and division within a place-value numbering system. It is mainly concerned with the underlying concepts and does not provide experience leading to assimilation of number 'bonds' - or seek to lead the child towards the meaningful application of the ideas to real life situations. Thus, M. A. B. experience is parelleled with other activities that lead to these ends.

"Again, the M. A. B. material introduces the child to a limited number of concepts. Others must begin to be abstracted from different experiences. Nevertheless, work with the M. A. B. is fairly extensive during the first year (age 7 - 8 years) in the Junior School, extends into the second year (age 8 - 9 years) for all children, and into the third year (age 9 - 10 years) for some slow learners.

"It is important to realize that the seven-year-olds work in varied number scales from the very beginning. No special significance is attached to our normal decimal scale until the child becomes aware of it for himself. Understanding is not hurried along by the teacher. Children work in pairs and develop individually in a permissive classroom 'climate.' The teacher must adopt an advisory role. There is no formal teaching. The M. A. B. is provided with an assignment card system that directs the children's activity, but teachers are encouraged to modify this as they grow more experienced.

"As one would imagine, few children get beyond addition and subtraction until they are nearing the age of 8. Division is tackled before multiplication because Dr. Dienes felt that the concepts involved were not so difficult, and children generally reach an understanding of these operations during their

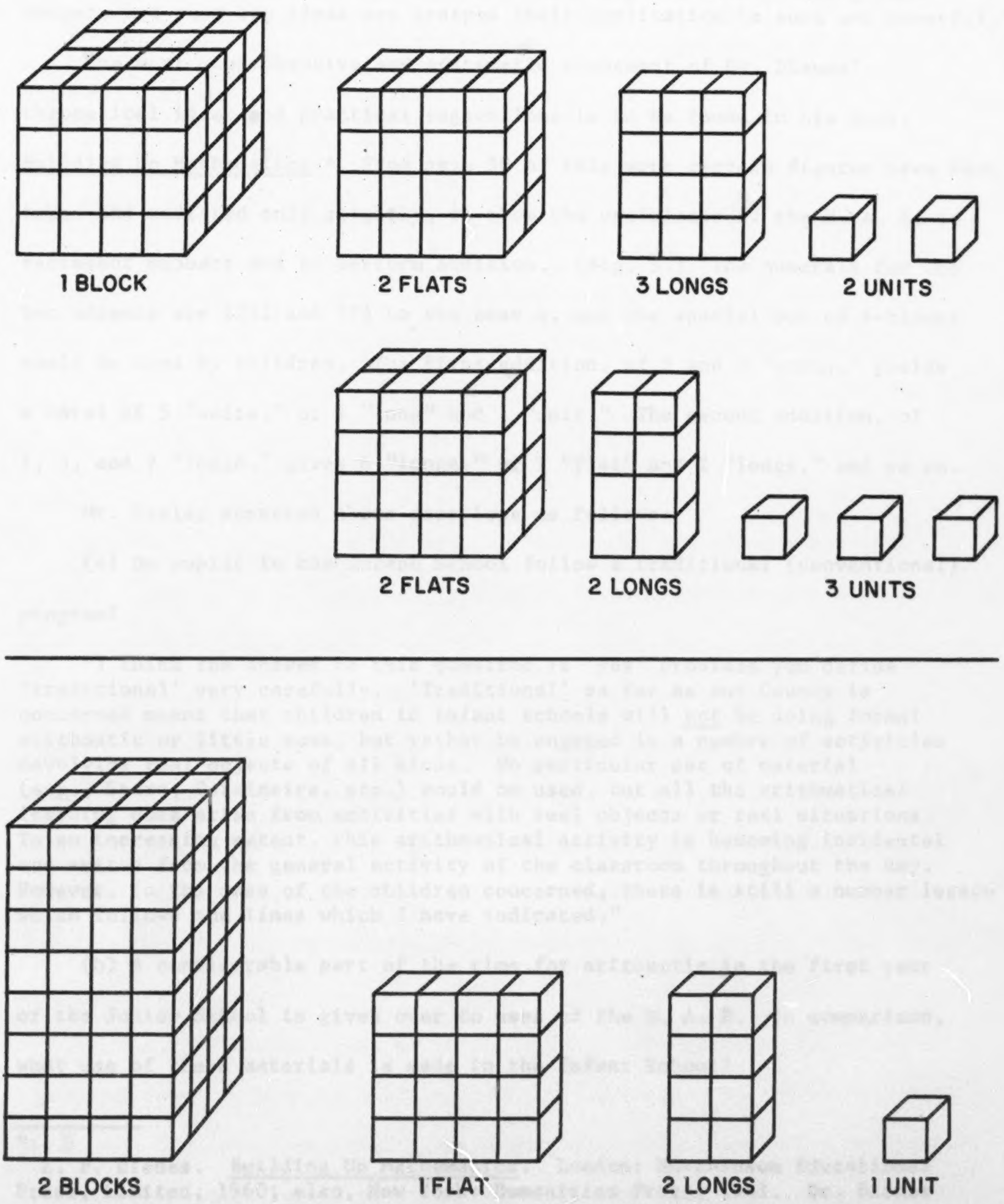
1 BLOCK
2 FLATS
3 LONGS
2 UNITS
2 FLATS
2 LONGS
3 UNITS
2 BLOCKS
1 FLAT
2 LONGS
1 UNIT

Fig. 5. Use of the M. A. B. to add 1232 (at the top) and 223 (next below), to get the sum, 2121 (at the bottom); base of 4. Adapted from p. 58 of: Dienes, Z. P. Building Up Mathematics, op cit. Putting together the 2 and the 3 "unit" blocks gives the total, 5 "unit" blocks, or 1 "long" and 1 "unit."

second year in the Junior School. Thus, children working with M. A. B. will not be able to calculate formally as soon as they could when the processes were taught in the traditional fashion. Learning and understanding take longer, but once the ideas are grasped their application is sure and powerful."

The most comprehensive and systematic treatment of Dr. Dienes' theoretical ideas and practical suggestions is to be found in his book, Building Up Mathematics.* From page 58 of this work certain figures have been taken and modified only slightly, to show the usefulness of the M. A. B. to represent numbers and to perform addition. (Fig. 5.) The numerals for the two addends are 1232 and 223 to the base 4, and the special box of 4-blocks would be used by children. The first addition, of 2 and 3 "units," yields a total of 5 "units," or 1 "long" and 1 "unit." The second addition, of 1, 3, and 2 "longs," gives 6 "longs," or 1 "flat" and 2 "longs," and so on.

Mr. Sealey answered three questions as follows:

(a) Do pupils in the Infant School follow a traditional (conventional) program?

"I think the answer to this question is 'yes' provided you define 'traditional' very carefully. 'Traditional' as far as our County is concerned means that children in infant schools will not be doing formal arithmetic or little sums, but rather be engaged in a number of activities involving real objects of all kinds. No particular set of material (e.g., Stern, Cuisinaire, etc.) would be used, but all the arithmetical learning does arise from activities with real objects or real situations. To an increasing extent, this arithmetical activity is becoming incidental and arises from the general activity of the classroom throughout the day. However, in the case of the children concerned, there is still a number lesson which follows the lines which I have indicated."

(b) A considerable part of the time for arithmetic in the first year of the Junior School is given over to uses of the M. A. B. In comparison, what use of these materials is made in the Infant School?

---

* Z. P. Dienes. Building Up Mathematics. London: Hutchinson Educational Press, Limited, 1960; also, New York: Humanities Press, 1961. Dr. Dienes earlier outlined the psychological basis of his system of instruction under the title, The Growth of Mathematical Concepts in Children Through Experience. London: Newness Educational Publishing Co. Ltd., 1959. This monograph is an offprint of Vol. II, No. 1, in the series, Educational Research, published by the National Foundation for Educational Research in England and Wales, London. Others of Dr. Dienes' statements appear in sources not readily accessible to the average reader.

"In the case of the --- School, the material is not introduced at all in the Infant School. In the case of the other three schools, the material is introduced to the children during the last term [third] in the Infant School, and they begin to play with it, engaging in building and other activities, with just a little stiffening of structured games when the child seems to be ready for them. No children work with the cards that are provided with the M. A. B. material, and the total experience is an introductory play stage."

(c) What part of the total time allotted to arithmetic in the first year of the Junior School is spent with the M. A. B.?

"One of the four cooperating Leicestershire Schools, the one which withheld the M. A. B. in the first two years, gave about 80% of the time to the M. A. B. in the first two terms, and 40% in the third term of the first year in the Junior School. In the other three schools, the per cents for the M. A. B. ran 40, 45, and 66."

The Conventional (Con.) Program

The Con. program is by all odds the oldest of the three considered in this report. For at least three decades it has been the most widely used system of instruction in arithmetic in the United Kingdom; and it is still the most popular. However, it has been under fire for some time, and it is being modified in many places and abandoned in others. Changes are in the direction of incorporating some of the theoretical features and/or of the methodology (including concrete aids) characteristic of what is commonly referred to in the United Kingdom as "structural arithmetic." And abandonment, when it occurs, is in favor of one or another of the forms structural arithmetic takes.

Unlike the Cui. program, which gives ratio and measurement the major place and counting a minor place in teaching arithmetic in the first school years, the Con. program, starting with experiences in enumerating collections of discrete objects, moves on to experiences in grouping and regrouping. Unlike the Dien. program which, once it gets fully under way, stresses the discovery and use of fundamental mathematical laws, the Con. program, at least until the last decade, has tended to omit these broad generalizations in teaching in the early years and to deal with them, if at all, much later on.

The statements in the foregoing paragraph are sweeping in character. They over-simplify, and they may misrepresent the extent of differences among the three programs under investigation. At this point, however, they have been made deliberately the way they read, in order to emphasize the fact of difference. In the concluding section of the chapter, after all three programs have been described by persons who know them intimately, there will be a place for greater nicety and greater thoroughness in comparison.

The statement regarding the Con. program was prepared by Mr. J. Clitheroe, M.A., Deputy Director of Education in the Carlisle, England, school authority. Mr. Clitheroe is a man of long experience in the administration of English schools and of discerning interest in problems of the curriculum and of instruction.

Mr. Clitheroe was a fortunate choice for an additional reason. Just before the start of the present investigation, an informal experiment in the teaching of arithmetic had been concluded in selected Carlisle schools. In the course of this three-year study Mr. Clitheroe had unusual opportunities to observe the activities engaged in by pupils and by teachers assigned to Cui. and to Con. schools. He also had the chance through conversation to learn directly from teachers their day-by-day reactions to what they were doing, whether favorable or otherwise. And in the final report he was able (a) to compare the two groups of schools for effectiveness of instruction as determined by paper-and-pencil tests and (b) to get from teachers their final evaluations of the programs they had taught as recorded in responses to a questionnaire.*

Mr. Clitheroe's statement follows:

"Before the last world war, it was both the official and the generally held belief in English schools that Arithmetic in the Primary School should mainly be concerned with the fundamental processes or rules. The Board of Education's Consultative Committee in its 1931 Report on 'The Primary School' stated that 'The chief problem for the teacher is how to secure a thorough mastery of these basic operations.'

"Emphasis was placed on skill in computation rather than on understanding and the ability to reason. Classes in Primary Schools and Departments were large. The legal "maximum" of fifty was far too often the common minimum. Teaching was done formally to the class as a whole. Each timetable showed a daily forty-five minute period for Arithmetic usually commencing about 9:30 or 10:00 a.m., the idea being that the young minds would then be fresh and at their most receptive.

"Each lesson tended to follow a similar pattern: a few opening oral questions to test skills already acquired, followed by five minutes of learning tables by Rote. 'Automatic availability' of knowledge was the goal. The daily chanting of tables was considered the best way of achieving this desired end. Not only multiplication tables were chanted, but tables illustrating the other three rules were also included. Normally, the rule itself was chanted; e.g., '6 and 5 make 11. Put down the one and carry one.'

"When the hypnotic five minutes of chanting were finished, the lesson would proceed, perhaps first with some exposition on the blackboard by the teacher; and then the whole class would spend the rest of the period in doing written work. All the children in the class would be attempting the same sums at the same time. It was not appreciated that this was often a waste of time for the bright, and a bewildering experience for the dull.

"The revolution came after the war and began in Infant Schools and classes. 'Teaching by telling' gave way to 'learning by doing.' Inevitably, the early enthusiasts went too far; and in many eyes if 'free activity' did not spell 'licence' it spelt 'anarchy' instead.

"Gradually the enthusiasm tempered into something controlled and directed. Woven into the children's activities were number games and experiences designed to ensure understanding of the rules. Playing at shops should lead to accuracy with real pennies! Playing with balances is a first introduction to weights; and measuring classrooms and other children's heights, to feet and inches.

---

* The report of this study has not been published. However, it can be said (a) that, though as a group they were superior to the Con. subjects in brightness and academic achievement, the Cui. subjects were little if any more competent arithmetically, as shown by their test scores, and (b) that the Cui. teachers ended the study in favor of the new program as compared with the Con. program which they had previously followed.

"Fundamentally, all these activities were based on the old, the traditional method of learning numbers by counting. The approach to learning had changed. The language remained the same.

"Junior School teachers had at first looked with disfavour on what was happening in the Infant Schools; but, as in general, 'activity' ceased to be completely 'free' and became controlled and purposeful, they realized they had much to learn from the Infant School colleagues about children's understanding and about the way they learn. Recognition of the individual needs of children has now won general acceptance, though in the Junior School there is still some sacrifice to the Gods of Speed and Accuracy as the children approach the end of their Primary life and their allocation to a Secondary School."

Mr. Clitheroe has described a sample of the materials and the activities used to develop concepts of the numbers, of the operations (almost entirely addition and subtraction in the first three years), and of the number combinations. He writes:

> "Of all the devices used to get children familiar with number concepts, the one most frequently used is the 'Domino.' This consists of counters or dots arranged like a domino. Commonly, children start with one counter placed at the top left-hand corner; the second is placed in the top right-hand corner, and so on until finally the fifth goes in the middle to complete the domino. The figures below illustrate the progression.

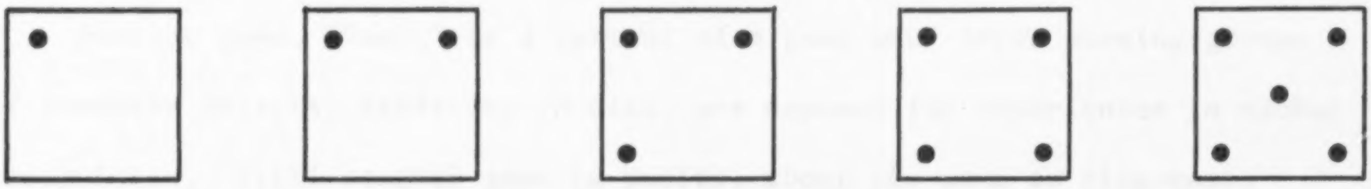

Fig. 6. Step-by-step development of the domino pattern for 5.

> "From this basic pattern all addition sums up to 5 are constructed - first by adding one's only, then sums involving the addition of two's, then three's. Similarly, the operation of 'take-away' is started by taking away one's, then two's, then three's.
>
> "Children advance to a concept of ten by a series of steps of 5 + 1, 5 + 2, etc., to 5 + 5 = 10. These addition bonds are then strengthened and extended by inverse drill on 'take-away' relationships: 10 - 1, 10 - 2, etc.

> "Counters, shells, and sticks are used in a somewhat similar fashion to make pictures, first, for one to five, then from five to ten; but children are encouraged to make with these materials patterns different from the domino; e.g., sticks in bundles, or in a long line, or in two lines, and so on; shells in small or large piles. By these means the teacher strives to convey to the child the idea that five, for example, is more than just five separate and distinct objects."

In some schools the writer saw various activities intended to teach the relative size of numbers by comparing and matching. One device involved facing pages in a booklet, the left page showing a regular grouping, say, for 8; the right page, a regular grouping, say, for 6. With markers, the children made the second group equal to the first by laying on the right page enough markers; and the discovered fact was then recorded, both as 6 + 2 = 8 and as 8 - 6 = 2.

Games are frequently employed; for example,

> "Skittles are used commonly up to nine. The skittles are set out in varying fashion, but nevertheless in some sort of group. The child starts by throwing one ball and registering how many he has knocked down, first, by making strokes on paper, and later by using numbers. Then the child throws two balls and makes a sum out of the two lots knocked down. Variations are introduced with two sets of skittles."

Another game, "Snap," is a variant of a game when cards showing groups of concrete objects, differing in size, are exposed for experiences in number recognition. Still another game is quoits, about the same as ring-toss.

Dice are employed in a progressive series of learning experiences.

> "To start with, the 'track' may be without figures [that is, only a line marked off in equal spaces], and the child records his progress by counting the spaces corresponding to the number on the thrown dice. Figures are entered for the spaces when the child is ready for them."

The social aim of arithmetic is taken care of in such ways as:

> learning the capacity of different vessels by measuring water in the low sink which is part of the classroom equipment

learning about money by shopping in the class shop; the children often make the shop and supply and price its articles

learning linear units of measurement, with accuracy in feet and inches, by measuring many things in and around the classroom

learning the concepts for shape -- round, square, diamond, triangle, oval, oblong -- by making patterns of the various shapes

learning a vocabulary of inexact quantitative terms, such as light, heavy, small, big, tall, short, and so on, by utilizing opportunities for relevant experiences as they arise throughout the school day.

It will have been noted that the description of the Con. program in British schools seems to fit the program most typical of American schools, certainly until recently. The two programs do have several features in common, chiefly in their basic theory. In the United States as in the United Kingdom the initial school experiences with numbers are with discrete objects, natural (beans, chairs) and conventionalized (markers, counters, cut-out squares), which are manipulated and enumerated. From these experiences it is hoped that the idea of numbers as groups will emerge (not the idea of numbers in terms of ratio, as in the Cui. program). Understandings of numbers as groups are made more abstract as real or pictured groups of objects are withdrawn, and activities in regrouping by combining and separating follow, leading to the relationships expressed in the number combinations in the four operations.

But the Con. program in British schools differs materially from its supposed American counterpart, even in the activities just described. From his classroom observations here and in the United Kingdom and from his examination of instructional materials the writer ventures the opinion that the American program is much more diversified in kinds of experiences provided children, puts greater emphasis on the development of mathematical meaning prior to encouraging memorization and computational skill, and stresses a much wider range of number relationships and mathematical principles. It is essential that such dissimilarities be made explicit at the outset, lest the reader assume that the findings with respect to the Con. program as taught in British schools are just about what would be obtained by an investigation of the American "conventional" program.

Comparisons and Warnings

(a) Variation vs uniformity of actual programs

Up to now it has been necessary to speak of the three programs as if each represented a single fixed, invariable system of instruction, -- as if, for example, in all Cui. schools teaching followed precisely the same pattern. Such was not at all the case, either for the Cui. schools or for the Dien. schools and the Con. schools. To mention one specific only, the time allotments for arithmetic were widely dissimilar among schools in the same program group.

Among Cui. schools there were differences in the degree to which reliance was placed exclusively on experience with rods as means of learning. Among Dien. schools, as Mr. Sealey has pointed out, there were differences in the time when the M. A. B. were introduced, and therefore in the proportion of total learning time given over to their use. Among Con. schools there were marked unlikenesses in the sequence of learning activities and in the material aids provided.

(b) Distinctiveness vs overlapping in program categories

Up to now it has been necessary, too, to speak of the programs as if they were completely, or almost completely distinct one from another. This implied uniqueness of each program is also contrary to fact. As might be expected, but for what may be called negative reasons, distinctiveness most nearly characterized the Con. program; but even so experimentation was under way here and there among Con. schools, to test the usefulness of materials and procedures not historically associated with this program.

In general, there was more or less overlapping between the three programs, and, in some instances, differences were matters of degree. In Dien. schools, counting, the basic operation in the Con. program, was not only permitted but encouraged in early stages of learning. In the Cui. schools, counting, certainly allowable in M. Cuisinaire's view but frowned upon in Dr. Gattegno's formulation, was certainly far commoner in classrooms in 1962 than it was in 1960 when the writer had last visited British schools.

The changes rapidly occurring in the Cui. program, most of them tending to increase the amount of overlapping with the Con. program, are perhaps best illustrated by the fact that in the United Kingdom the writer was unable to recruit a single "pure" or "doctrinaire" Cui. school. Local administrators were regularly requested to nominate schools of this kind; but none was forthcoming. In one of the largest cooperating Scottish school authorities, where the Cui. program had been officially adopted for general use, only three of the many schools (90 pupils) were made available, -- and this, with the comment that these schools "probably" came "closest" to meeting the criterion. Other schools in the authority were suspects of too many practices smacking of the Con. program.*

(c) Real differences among programs

In view of the facts set forth in (a) and (b) above, how is it possible to compare the results of instruction according to "the three programs," in promoting maturity of arithmetical conceptualization (or anything else)? The answer to the question, and the justification for the comparisons to be made, is that, despite variations of particular programs within program categories and despite overlapping between the main program categories, there are differences in the three programs in many important respects.

---

* No adverse criticism is intended. The perfect or best way of teaching arithmetic has not yet been found. Until it is ready at hand, experimentation is to be encouraged, and with all sorts of compromise and amalgamation of supposedly irreconcilable approaches.

(1) No children in Con. and Dien. schools had had experiences with the Cui. rods, or had even seen them. They therefore were subject to differences not only in material aids, but also in arithmetical objectives, as below.

(2) No children in Con. or Cui. schools had had experiences with the M. A. B., or had even seen them, the consequence being differences of the kinds in (1).

(3) In Con. schools, but not in Cui. and in Dien. schools, the typical arithmetical objectives are: early mastery of the number facts, competence in computation as far as computation is carried, and ability to solve problems of the formal verbal type.

(4) Emphasis on "social arithmetic" and on informal experiences with actual quantitative situations was commoner in Dien. schools than in schools that had adopted either of the two other programs.

(5) Instruction was generally most formal and routinized in Con. schools; less so in Cui. schools; least of all in Dien. schools.

(6) Activities involving ratio and measurement as the basic means of teaching arithmetic were stressed much more in Cui. schools than in Con. and Dien. schools.

(7) In the matter of over-riding mathematical principles and laws, those selected for emphasis differed in Cui. and in Dien. schools, but

regardless of kind they were relatively neglected in Con. schools.

(d) Pre-judgment as related to arithmetical tasks

According to the analyses in (c), the concepts and skills selected for interviews may seem to have been given an unfair advantage to the Con. subjects over children taught according to the Cui. and the Dien. programs. This inference is valid only in part.

(1) Number facts: early mastery of the facts is stressed more in Con. schools than in other types of school; hence the pupils in Con. schools might be expected to outscore those in other types of school in rate and accuracy. On the other hand, level of conceptual development is something quite different from proficiency, and the Cui. and Dien. children had at least as good a chance as had the Con. children, to earn high ratings in this respect.

(2) Computation: again in measures of rate and accuracy, the Con. subjects might be expected to score highest, for emphasis on skill in computation is deferred in Dien. and Cui. schools. But the center of interest in this study is level of conceptual maturity, and Cui. and Dien. children may well have developed understanding of the mathematical rationale of computation in relatively greater degrees.

(3) Problem solving: There is no way of knowing which of the two groups of children, if any, had had more experience in solving formal verbal problems; hence, the Con. subjects might or might not have been expected to score higher than others in proficiency. On the other hand, it is to be noted that effort was exercised in the interviews to make the problems natural and realistic. And as for quality of reasons given for the chosen procedures, the Cui. and the Dien. children may have had experiences that put the Con. children at a disadvantage; that is to say, the former groups may have been

better taught the functions of the four operations and called upon oftener to verbalize their reasons.

(e) Re-statement of the bases of comparison

The re-listing of the three types of arithmetical task included in the interviews should serve to warn the reader of many aspects of arithmetic which were not included. No chance was given the Cui. children to reveal unmistakably and comprehensively the effects of their familiarity with the ideas of ratio and measurement, or their grasp of other mathematical relationships learned through the use of the rods. Nor was any chance given the Dien. subjects to disclose their ability to employ the mathematical principles and generalizations peculiarly a part of the Dien. program. And, as a last example, no chance was given to any group at all to show how well it could deal with the quantitative situations of life in their natural settings.

In a word, the base for comparing the results of instruction by the three programs is a narrow one, and the arithmetical tasks selected to get comparative data are few in number. Nevertheless, it is believed that the basis of comparison is broad enough, and that the means taken to secure data are numerous enough to produce adequate evidence relevant to the major problem; namely, how do the three programs compare in their effectiveness in fostering progress toward arithmetical abstractions and toward mature arithmetical conceptualization?

CHAPTER 3

DETAILS OF THE PROCEDURES

The Training of Interviewers

Typically, interviewers were trained in groups of three or four in two full days. For fewer or more interviewers, the period was shortened or lengthened. What happened in a two-day training period is outlined below.

In the first morning session the nature and purpose of the research were described; reasons were given for the use of the interview rather than paper-and-pencil testing; interview schedules were examined, and categories and their corresponding code symbols were explained. Noting the complexity of the schedules -- about ninety separate items had to be entered for each child -- usually in each group one or more interviewers expressed doubt of ever being able to learn what they were to do, and at least one interviewer requested a list of the specific questions to be asked of children. (By the end of the training period, doubts had just about disappeared, and the impracticability of set questions had become self-evident.)

Next, a child of appropriate age and schooling was brought into the conference room and was seated at a table facing the writer, with his back to the interviewers. The writer then demonstrated the interview, explaining as he went along how and why he classified each response as he did. At the conclusion of the interview, there was general discussion of the questioning and of the classification of responses.

A second child was then interviewed by the writer. This time the latter did not announce his classifications, but asked the interviewers to do so, each for himself. When the child had left, classifications were compared, and discrepancies were discussed with a view to securing agreement. Queries about the relevancy and purpose of the interviewer's questioning were encouraged.

In the afternoon session, one of the interviewers was asked to try his hand with a third child. The actual interviewer and the observing interviewers entered responses and classifications independently. During the course of the interview, the writer interrupted from time to time to criticize and to illustrate more fruitful forms of questioning. When the interviewer seemed to be stymied, the writer took over temporarily. At the completion of the interview there was the usual period of comparison of entries, of discussion, and of comment.

Thereafter for that day and the next, children were questioned by interviewers, one after another, always under the observation of the writer and of the other interviewers. Upon request, he questioned still other children at points where the trainees felt most insecure. All in all, each interviewer worked with at least three children himself, if not on the whole schedule, then in one instance on that part of the schedule where he needed most help. In addition, each interviewer observed two complete and some incomplete interviews by the writer as well as at least four full and some partial interviews conducted by his colleagues.

The subjects in the practice interviews were seldom included in the research sample, -- only when in the writer's opinion their responses were in no way made invalid by happenings in such interviews. Nor were other subjects included when they were used in privately arranged sessions, in order to permit interviewers to improve their technique of questioning children by practicing in the presence of other trainees who were asked to criticize.

In the course of preparing themselves to collect research data, the interviewers had ample opportunity to learn to observe the following cautions:

(a) to avoid cues and leading questions from which children could profit;

(b) to refuse to stop questioning when only superficial answers were given, for such answers could not reveal what was wanted;

(c) to adapt the pace of interviewing to the pace at which each child could respond comfortably;

(d) to vary questions when responses were ambiguous, in order to draw children out;

(e) to maintain a pleasant, friendly, warm atmosphere;

(f) to stop interviews (they lasted about an hour on the average) as soon as any child showed rebelliousness, or boredom, or anxiety, bringing him back to complete the interview at another time if this was possible; otherwise, rejecting him as a subject; and

(g) to record exactly children's answers when the responses were especially interesting or resisted categorization.

Data and Method of Collection

Data on pupils

Interviewers asked each child his name and recorded it and his sex on the interview schedule. From the headmaster or headmistress he learned the number of minutes per week devoted to arithmetic in each school in the current year and in each of the two preceding years. He entered this information also on interview schedules. From the teacher or from the central office he obtained each child's age, recording it in one of four age categories, from 7 years 5 months or less, to 8 years 6 months or more.

There being no test scores available, interviewers asked teachers to rate each child for brightness and for achievement in reading and in arithmetic. The estimates were made on a 4-point scale. Teachers were told not to distribute their pupils equally among the four quarters, but to think of them in terms of all the pupils they had had in their teaching experience. In the case of brightness, children were to be put in the top quarter (rating of 1) only if, in their teacher's opinion, they could succeed in the grammar school. Thus, one teacher's pupils might be classified for brightness like this: 0% rated 1, 28% rated 2, 45% rated 3, and 27% rated 4, whereas the per cents for another teacher might be 18, 50, 23, and 9, respectively. Corresponding directions were given for placing children according to academic achievement.

Interview data

1. The number combinations. Data were procured on the number combinations in three steps.

(a) A child was first given a slip of paper on which had been printed the easy combinations

$$\begin{array}{r}3\\\underline{+\ 4}\end{array},\quad \begin{array}{r}6\\\underline{-\ 2}\end{array},\quad \begin{array}{r}2\\\underline{\times\ 3}\end{array},\quad 6\overline{)12}$$

He was then asked to read the combinations and give the answers, one by one. The purpose was to find and record the way he thought of the combinations in the four operations so that the same language forms could be used with the eight interview combinations. Thus, a child who read $\begin{array}{r}3\\\underline{+4}\end{array}$ as "Four plus three are seven" was asked later, for $\begin{array}{r}8\\\underline{+5}\end{array}$, "How many are five plus eight?," and not "How many are eight, add five?," or "How many are eight and five?"

(b) Next, each child was given orally the eight interview combinations. At this time, only two entries, in addition to the answer given, were made for each combination: for accuracy (Correct Answers, Incorrect Answers, or Non-Attempt) and for speed (Prompt Response, Slight Pause, or Long Pause). (No entry for speed was recorded when no attempt was made.) Ratings for speed were estimated in terms of what appeared to be the child's normal quickness of response. Thus, original ratings of Slight Pause were often changed to Prompt Response when it later became evident that for him he had responded at his customary rate.

(c) As the last step in the entire interview, the same eight number combinations were again presented orally, but for a different purpose. This time the interviewer elicited from the child his method of thinking of each combination , this method to be classified according to type, of which there were twenty. The two categories Non-Attempt and Indeterminate are self-explanatory. Many of the other categories will need to be described, and this will be done shortly.

Questions relating to thought processes were deliberately avoided the first time through with the number combinations, so as not to get pupils concerned too soon about such matters. Questions designed for this purpose, coming as they did at the end of the interview, did not disturb them, probably because of intervening activities. Not more than one child in fifty gave convincing evidence that he was aware of having responded to the combinations at an earlier time.

2. Computation. The four interview examples at the right were printed on a slip of paper which was placed before a child. He was instructed to find the answers, talking aloud at the same time and writing whatever he wanted to. He was not interrupted until after he had attempted the last example. Meanwhile, the interviewer had entered on his schedule Correct, Incorrect, or No

$$\begin{array}{r} 28 \\ \times\ 3 \\ \hline \end{array}$$

$$\begin{array}{r} 92 \\ -\ 45 \\ \hline \end{array}$$

$$4\overline{)84}$$

$$\begin{array}{r} 27 \\ +\ 46 \\ \hline \end{array}$$

Attempt, as measures of accuracy, and Smooth or Erratic as an estimate of level of performance. He had also copied on the blank whatever the child had written on his paper.

After the child had completed the fourth example, he was taken back to the first example, then to the second, and so on. In each instance, he was questioned to see in how far he understood the mathematical principles of computation in what he had done. (Such questioning was omitted for examples which were not attempted.) Explanations were classified as Satisfactory or as Unsatisfactory, each with the sub-categories, Words, Altered Examples, Images, and Other. The meaning of these terms will be made clear at a later point.

3. Verbal problems. Holding before himself a sheet of paper containing the twelve problems, the interviewer presented each problem orally and informally, seldom in the set, abbreviated form in which it appeared on the paper. Often the interviewer turned a problem into a story, and always he insured understanding of the problem setting and knowledge of the numbers before calling for an answer.

For example, an interviewer might say, "Do you like eggs? ... Did you ever collect eggs in a henhouse? ... Well, this story is about a little boy named Sam. One day his mother asked him to go to the henhouse and get all the eggs he could find. When he was through, he had sixteen eggs in his basket. How many eggs? ... Yes, sixteen eggs. On the way to the house, he slipped and fell. What do you suppose happened? ... Yes, he broke some of the eggs. When he counted them, he found that seven eggs were cracked. How many eggs were cracked, ... Yes, seven. And how many eggs were in the basket? ... Yes, sixteen. Now tell me, how many eggs were not cracked?"

Accuracy data were recorded as Correct Answer, Incorrect Answer, and Non-

Attempts. Data on level of performance were entered as Low, Intermediate, and High based upon observations of the way the child went about his solutions. Finally, the child was questioned in order to ascertain how well he could account for his method of attack. (Entries for level of performance and quality of reasoning were made only for attempted solutions.) Explanations of reasons were rated as Good, Fair, Unsatisfactory, and Indeterminate. The last-named category was used only when a child's verbalization, seemingly with some degree of understanding on his part, was nevertheless unintelligible to the interviewer.

Order of record-making. In summary, data for each interviewee were obtained in the following series of steps:

(1.) Personal data: name and sex of interviewee.

(2.) The number combinations: proficiency data only.

(3.) Computation: proficiency data first; then immediately data on knowledge of mathematical principles involved in computation.

(4.) Verbal problems: data on proficiency, then on quality of reasoning.

(5.) Number combinations a second time: data on thought processes.

(6.) Data from central office: time allotment for arithmetic. Data from teacher: child's age, and estimates of his relative brightness, reading achievement, and arithmetic achievement.

Classification for Conceptual Maturity

Data on proficiency -- accuracy and rate of work or level of performance -- were readily obtainable by observing what a child did and said voluntarily; but not so in the case of data on thought processes, quality of understanding, and the like. The technique for securing the last-named types of information was the interview. The procedure used for each type of arithmetic task will next be described. These descriptions, necessarily much shorter than is desirable, may give the impression of abruptness on the part of interviewers and of lack of consideration for interviewees. It needs to be stated emphatically that children were never "pilloried" or "pursued," but were always treated with respect, patience, and sympathy. The nature and the rate of questioning were adapted as needed for each child.

Computation

Data on his proficiency in computation having been entered on the schedule, the interviewee was taken back over the four examples. This time interest centered in his ability to explain what he had done in terms of place value or of renaming numbers.

Suppose that a child in dealing with the multiplication example, wrote his work as at the right, meanwhile saying, "Three times eight, twenty-four. Three times two, six; and two more are eight." The interviewer then asked him to explain what he had done. If he was able to do so, he was not questioned on this example, but was taken on to the next.

$$\begin{array}{r} 28 \\ \times\ 3 \\ \hline 84 \\ 2 \end{array}$$

If, however, questioning was required, it might go something like this: (a) "You said, 'Three times eight are twenty-four,' didn't you?" ... (b) "Well, under the 8 and the 3, I see 4, not 24. Why is this?" ... (c) "Could you have written all of the 24 under the 8 and the 3? Why, or why not?" ... (d) "What is this little 2 I see under the 8 in the answer? Where did you get it?" ... (e) "What does it mean? Why did you write it where you did?" ... (f) "Did you use this little 2 at all? If so, how?" ... (g) "Why could you use the little 2 in this way?"

The response of a child who satisfactorily answered questions (a) to (g), or comparable questions, was classified as Satisfactory, Altered Example, -- Satisfactory because he knew what he was doing and could explain it; Altered Example because he had changed the physical appearance of the example, to get the advantage of a crutch he needed. Satisfactory, Words Only, was the category name for responses made by children who did not write the crutch and could give their explanations without the use of any external aid of this kind; for them, words were enough.

Other types of responses were also classified under the general rubric, Satisfactory; namely, (a) explanations given verbally but admittedly mediated by some kind of mental imagery, such as images of counters of Cuisinaire rods; and (b) explanations which were entirely sensible mathematically but did not involve explicit reference to place value and renaming. As an example of (b), a child's procedure for 3 x 28 might be, "Three thirties are ninety, less six; eighty-four," and the child using it might make it unmistakably clear that he knew what he was about. In this case, the entry was Satisfactory, Other, while that for the use of imagery was Satisfactory, Image.

There were therefore four entries under Satisfactory: Words Only, Altered Example, Image, and Other; and the same four sub-classes were used in categorizing responses which were judged to be Unsatisfactory. Thus, a child, when asked to explain his computation, could merely repeat what he had said while at work, with no real explanation at all. If he did so without writing a crutch, his response was scored as Unsatisfactory, Words Only; if he did use a crutch, as Unsatisfactory, Altered Example. Similarly, a child could verbalize meaninglessly if he employed images (Unsatisfactory, Image) or some kind of computational procedure not readily verbalized in terms of the renaming of numbers (Unsatisfactory, Other).

Problems

Children's explanations of, and reasons for their solutions of problems were sometimes evoked only with difficulty, and then might not be susceptible to confident classification. In some instances, obstacles seemed to arise from the unusualness of the requirement to articulate the thought processes used and to supply reasons therefor: that is, they did not know what to say. Perhaps they lacked the technical terminology that was needed, though there was no insistence upon the use of such words as "add" or "product" (and when not used by children they were not supplied), but the greater probability appeared to be their lack of experience in answering the questions beginning with How, and Why. The best that could be done in such cases was to record responses as Unsatisfactory, when a child rambled at random, or Indeterminate, when he appeared to have an inkling of the relevant reason and could talk about it, even if not in a manner completely intelligible to the interviewer.

Fortunately, the majority of explanations were subject to classification as Fair or Good. The top rating, Good, was assigned only when children knew unmistakably what they had done, and why, and could verbalize explanations and reasons in a sensible way and in a way which related selected operations to their true mathematical functions. The lower rating, Fair*, was given responses inferior to Good, but much better than Indeterminate (in that they were appropriate to the given mathematical situation and could be readily understood), and much better than Unsatisfactory.

---

* For reasons which will appear later, per cents of explanations classified as Fair had to be eliminated in comparisons.

The method of questioning may be illustrated by what could have been done with a hypothetical child who correctly obtained the answer, 4, for the following problem:

> Susan had a box of 12 crayons. She can find only 8 of them.
> How many of her 12 crayons are missing?

(a) "How did you get 4 for your answer? That is, what did you do to get 4?"

(b) "What made you think that that was the way to find the answer?", or, alternatively, if the child had himself introduced the word "subtract," "What made you think that you should subtract?"

(c) "Could the answer be 3 (or 7, or 16, or whatever)? Why, or why not?"

It is impossible to list more than these three basic questions, for the interview might be given a new direction by the answer to any question. For example, the child might defend his correct answer, 4, by saying that he had to subtract in order to have "a smaller answer." In such an instance, he might then be asked, "Well, you could get a smaller number by dividing 12 by 8, couldn't you? Why didn't you do that?" Or, another child who subtracted upon hearing the (to him) cue word "missing," might be questioned to see how long, and why, he would hold to this reason.

Question (c) above, which provides incorrect answers, was asked to check a child's confidence in his answer and sometimes to elicit by negative suggestions explanations that were not forthcoming in response to the more positive questions (a) and (b).

It needs only to be added that much the same kind of questions were used when children gave incorrect answers. Their reasoning might be sound and straight-forward, their errors being due to slips of memory or to having learned wrong answers to number combinations.

Number combinations

The final list of categories of mental processes used with number combinations, made up after five revisions, totalled twenty. Eighteen of these were derived from tryouts with American and English children, and two were added as logical possibilities.

The category Non-Attempt is self-explanatory. Indeterminate was recorded when interviewers were unable to evoke intelligible verbalizations.

Six categories were set up for counting, three each for Counting by 1's and for Partial Counting. Three sub-categories both for counting by 1's and for partial counting were intended to discriminate among the media employed: Objects, Images, and Words. In both counting by 1's and partial counting children could use as objects such things as finger movements, tapping, marks made on paper; as images, "mental pictures" of markers, rods, formalized drawings in workbooks; as words, just the numerals in the number sequence.

Counting by 1's was recorded when children, using objects, images, or words alone, started with 1 and reported their procedure as follows for 8 + 5, "1, 2, 3, 4, ... 13"; for 15 - 7, "15, 14, 13, ... 8"; for 6 x 4, "1, 2, 3, 4; 5, 6, 7, 8; 9, 10, 11, 12," and so on, making up six collections of four each, or four collections of six each; for $7/\overline{21}$, "21, 20, 19, 18, 17, 16, 15, 14; 13, 12, ... 7; 6, 5, ... 0," or anything similar.

Partial Counting did not occur in the case of multiplication and division combinations. It did occur, with objects, or images, or words, with addition combinations (e.g., for 8 + 5, "8, 9, 10, ... 13," ... or "9, 10, 11, 12, 13," or "6, 7, 8, ... 13") and with subtraction combinations (e.g., 15 - 7, "8, 9, 10, ...

15").*

Recitation of Tables covered such instances as when children, for 6 x 4, started, "1 x 4 = 4; 2 x 4 = 8," and continued to 6 x 4 = 24," or began, say, with "4 x 4 = 16" and went on from there. In other words, they did not know 6 x 4 = 24 as an independent fact, but found it necessary to build up to it. This category covered also corresponding procedures with the two division combinations, but very rarely appeared in the case of addition and subtraction combinations.

Imagery was the category for responses when children thought of number combinations in terms of imagined markers, rods, and the like, but without counting or partial counting.

Fairly often children found an answer for a combination by using its reverse; thus, 15 - 7 = 8 because 15 - 8 = 7; 6 x 4 = 24 because 4 x 6 = 24; 8 + 5 = 13 because 5 + 8 = 13; 21 ÷ 7 = 3 because 21 ÷ 3 = 7. Since this kind of procedure could be employed with or without understanding of the reason why the practice was effective, two categories were devised, Reverse, Rational, and Reverse, Mechanical. Differentiation was made by questioning.

Inverses also were quite common; thus, 15 - 7 = 8 because 7 + 8 = 15, or 8 + 7 = 15; 6 x 4 = 24 because 24 ÷ 6 = 4, or 24 ÷ 4 = 6; 8 + 5 = 13 because 13 - 5 = 8, or 13 - 8 = 5; 21 ÷ 7 = 3 because 3 x 7 = 21, or 7 x 3 = 21. For convenience, procedures like the series 24 - 6 = 18; 18 - 6 = 12; 12 - 6 = 6, and 6 - 6 = 0, were classed as inverses for 6 x 4, as were procedures like

---

* Note that, in subtraction, procedures involving counting backward from the sum were classified as Counting by 1's, since in counting backward a child had to start with the sum. On the other hand, in subtraction when a child counted forward as in the illustration above, his response was classified as Partial Counting: he could have started with 1. The validity of this differentiation is arguable; but, as will appear, the number of cases involved is too small to affect interpretation of the data.

7 + 7 + 7 = 21 for 21 ÷ 7. Again two categories were needed, Inverses, Rational, and Inverses, Mechanical.

For clasification under one or the other of the two last-named categories, a child had to start and to "carry through" with the inverse to get his answer. There were instances when initial use of an inverse was completed in another way (but without counting); e.g., for 32 ÷ 8, the method of thinking might be 2 x 8 = 16, and another 8 is 24, and another 8 is 32; so 32 ÷ 8 = 4. Such procedures were classed as Inverses Plus, a form of rational solution.

Four other categories were Rote Recall, Guessing, Solution, and Meaningful Habituation, all of which, as far as the child's initial responses are concerned, may seem to be the same. Careful questioning, without clues of any kind, enables one to discriminate one from the others, however.

When a child insisted upon the correctness of his answer, even when it was incorrect, repeating it over and over again and resisting suggestions of other answers, but was unable to justify his answer or to "prove" its accuracy, the entry was Rote Recall. When, however, a child changed his answers, being apparently as well satisfied with one as with another, the entry was Guessing.

When a child reported that he had first thought something like 8 + 2 = 10, + 3 = 13 for 5 + 8, before saying "Thirteen," his response was recorded as a Solution, as was a response like 17 - 7 = 10, - 1 = 9 for 17 - 8. The essence of the category Solution is that the child does not know the answer, but must and can find it, and that he does so by using number relationships of a higher order than counting. The term Solution is none too satisfactory for it has to be defined, for present purposes, as categorizing thought processes more advanced than counting and excluding certain others of the kinds of process described above which partake of the nature of "solution."

The category, Meaningful Habituation, was reserved for responses that were, first of all, prompt and correct. (Rote Recall responses could have both characteristics, as could Guessing occasionally; but Solution responses were usually a bit slow.) Responses were classed as representing the use of Meaningful Habituations when the child, if asked, could demonstrate his answer to be correct by using one or another form of Solution. It is to be noted that he did not use Solution to get his answer, which he furnished at once. But by citing "proof" of his announced answer later when requested to do so, he showed that he had real and functional understanding of the relationships involved, these supplying a kind of background for his immediate confident answer.

The twentieth category, Other, was listed on the chance that some responses, other than those classed as Indeterminate, might occur, in which case it would have been necessary to recognize and name a new class or classes. None, however, was reported. The category, Other, then was one of the two needlessly included for logical reasons, the other being Partial Counting, Images.

Treatment of the Data

The data will be presented in tabular form more completely in Chapter 4, and the discussion will be fuller, than in later chapters. Once the reader has worked through Chapter 4, which illustrates the pattern of analysis to be used generally and the types of comparison to be made, he will be able to understand and interpret the shorter and less detailed method of treating the data thereafter. Beginning with Chapter 5, some tables like those in Chapter 4 will be omitted, and in others some categories will be eliminated when they serve no purpose, or will be combined with other categories of like import.

The data will be dealt with in a very simple manner. Nearly all measures will be expressed as per cents or as average per cents, and comparisons will be made in these terms.

It is hoped that the explanations earlier in this chapter will have established the face validity of the data. As for their reliability, the reader will recall the statement that, after training, interviewers were in close agreement in classifying children's responses, and the training period was extended as much as necessary in order to assure that this would be so.

As is well known, a per cent of 0.06 and a percentage difference of 0.36 may be "significant" statistically (that is, reliable), provided that the group or groups involved are large enough and the variability of measures is great enough. On the other hand, such small per cents mean little or nothing in problems relating to teaching and learning in the school. One would hesitate to say that an arithmetic program which produces an average of 91.7% in accuracy of computations is preferable to another that produces an average of 92.5%, no matter how reliable (i.e., statistically "significant") the difference may be. And, similarly the occurrence of a particular way of thinking of number combinations amounting to 3% of all attempts is hardly a matter of much consequence in the classroom.

Still, somehow the problem of "significance," whether viewed statistically or educationally, has to be resolved. In the present study, while the reliability of no per cent and of no difference between per cents has been computed by the usual statistical methods, "significance" in the sense of reliability is not neglected. Rather, it is made to turn upon arbitrary subjective decisions, with which equally arbitrary decisions concerning educational "significance" are combined. No difference between per cents will be said to be "significant" purely in the context of educational practice; it must first meet the subjective test of reliability. And an intermediate area between "significance" and "insignificance" of differences will be provided, in which differences are not described in these terms at all. By this method, it is hoped, the results of comparisons will always be evaluated in terms both of educationally and of statistically "significant" differences.

The following rules have been employed to interpret differences between per cents and between average per cents. (a) Differences of 10 points or more will be regarded as significant. They are almost certainly reliable, and they may well have important implications for teaching and learning. Differences larger than 12 or 14 points may be described by some such term as "considerable" and "substantial." (b) Differences of 5 or fewer points will be regarded as insignificant, for their reliability is doubtful and, on this account and others, they probably mean little if anything educationally. Such differences will rarely be reported. (c) Differences of between 6 and 9 points define the intermediate area referred to above, and they are viewed, at least at the lower end of this range, as being of doubtful reliability. They will be cited at times when they are of special interest, e.g., as being in line with, or directly opposed to the general body of evidence. When mentioned at all, differences of 5 to 9 points referred to with some such limiting term as "somewhat," or "possibly," or "slightly," or "perhaps larger (or smaller)."

Imperfect as is the scheme just outlined, the rules can be applied uniformly and impartially to the results obtained with the three instructional programs to be compared, with injustice to none. There is no denying that the decisions to be made represent the exercise of judgment. On the other hand, judgment continues, as it must, to play a major role in determining educational procedures, and it is by no means absent in all research, whether or not its presence is recognized and explicitly noted.

PART II

THE INVESTIGATION IN SCOTTISH SCHOOLS

CHAPTER 4

GROSS COMPARISON OF THE EFFECTIVENESS OF THE CUI. AND THE CON. PROGRAMS

Data on Schools and Subjects

The comparability of the total samples of subjects taught in accordance with the Cui. program and with the Con. program needs first to be examined.

In each school authority, only the schools known to be teaching the Cui. program most scrupulously were to be selected. There seemed to be no point in drawing subjects from nominally Cui. schools where this condition was not met, -- and there were many of them. In one of the largest cooperating authorities, where the Cui. program had been officially adopted some years before, the administrators were willing to nominate only three schools (90 subjects) for use in the study. Other schools in the authority were excluded, for in them it was believed features of the Con. program were still more or less dominating instruction. The probability is that administrators who chose only their "best" Cui. schools were equally careful to choose only their "best" Con. schools. Certainly they were urged to do so.

According to Table 1, subjects were supplied by seven Cui. schools and by six Con. schools, though in the end there were 28 fewer Cui. subjects. Variation in the numbers of subjects furnished by schools was large. For some reason one Con. school sent only five children to interviewers. Schools, even those within the same authority, differed greatly in the amount of time allowed for arithmetic. For example, in one school the weekly time allotment was 150 minutes in the first year, 275 in the second, and 250 in the third. In another school in the same authority the corresponding totals were 135, 150, and 225 minutes, respectively, by years. In some

Table 1. General Information
Concerning Cooperating Schools

Type of Program	Number of Schools	Total Number of Subjects	Range in Number of Subjects	Time Allotment for Arithmetic, 3-Year Total	
				Range	Weighted Mean*
Cui.	7	225	21 - 42	480 - 750	587
Con.	6	253	5 - 73	660 - 759	721

* See text.

schools, too, the largest amount of time devoted to arithmetic occurred in the first year and the smallest amount in the third year, whereas in other schools the time allotment was smallest in the first year and largest in the third. While it would be interesting and possibly valuable to explore the effects of different patterns of time allotment, the small number of schools in the various types of pattern precludes the possibility.

Instead, time allotment must be treated in gross terms, and for this purpose weighted means** were used. The weighted mean for total instructional time in Cui. schools is 587; in Con. schools, 721. However crude the method of computation, the weighted means are sufficiently trustworthy to indicate that, as compared with the typical Cui. subject, the typical Con. subject

---

** The weighted means for the Cui. and the Con. schools were calculated in the following manner. Assume that in School A, a Cui. school, the weekly allotments were 200, 300, and 250 minutes in the successive years, to give a total of 750 minutes, admittedly an artificial figure which cannot really be named. Next, multiply 750 by 25 (the number of subjects), and the total is 18,750. Then add the 25 (subjects) to the corresponding numbers of subjects in other Cui. schools, and add the 18,750 to the corresponding measures of total time for these schools. Finally, divide this last sum (time) by the total number of Cui. subjects, to get the average amount of time spent by Cui. subjects in the three years combined, the weighted mean.

was under instruction in arithmetic for a significantly larger amount of time, perhaps 20% more in the three-year span.

The data in Table 2 represent the actual per cents of Cui. and of Con. subjects with respect to five categories of characteristic. Probably the differences in the proportions of the two sexes are of little consequence educationally. The weighted means for the other four characteristics in Table 3 were calculated from the actual per cents in Table 2 in the manner already described for time allotment.

As can be noted in Tables 2 and 3, the Cui. and the Con. subjects differed, as groups, in a number of respects other than instructional program and amount of teaching time. For example, the Con. subjects have a higher rating for brightness than do the Cui. subjects.

Such differences will be temporarily neglected, in order to simplify the comparisons of the Cui. and the Con. groups with regard to various aspects of learning. Once these comparisons have been completed, the differences in question will be explicitly noted and summarized, and their possible effects on the findings will be examined. This will be done in the Chapter Summary.

The Number Combinations

The per cents in Tables 4 and 5 and in all similar tables are averages. In their calculation, decimals from 0.1 to 0.5 have been disregarded, and decimals from 0.6 to 0.9 have been rounded to 1.0.

Conceptual maturity: Level of thought processes

Level of conceptual maturity in the case of the number combinations was determined from the ways in which children thought answers.

Table 2. Comparison of Characteristics of Cui. Subjects and Con. Subjects, in Per Cents

Characteristic		Per Cents, by Program	
		Cui.	Con.
Sex	Male	47	53
	Female	53	47
Age	7 yr. 5 mo. or less	4	1
	7 yr. 6 mo. to 7 yr. 11 mo.	51	43
	8 yr. 0 mo. to 8 yr. 5 mo.	44	43
	8 yr. 6 mo. and over	2	13
Estimated Brightness	Rating 1. (Brightest)	20	26
	2.	27	28
	3.	29	30
	4.	24	16
Estimated General Achievement in Arithmetic	Rating 1. (Highest)	19	29
	2.	24	22
	3.	32	30
	4.	25	19
Estimated Reading Achievement	Rating 1. (Highest)	21	30
	2.	32	30
	3.	24	25
	4.	23	16

Table 3. Characteristics of Subjects, Weighted Means

Type of Program	Age	Rating in Estimates for		
		Brightness	Achievement in Arithmetic	Achievement in Reading
Cui.	7 yr. 9 mo.	2.57	2.46	2.63
Con.	7 yr. 10 mo.	2.35	2.26	2.39

All combinations. Each of the 225 Cui. subjects was asked to give answers to eight combinations, presented orally one after the other, thus making a total of 1,800 responses. The thought procedure used in finding each answer was then classed according to the scheme of twenty categories already described. The corresponding total of Con. responses was 2,024. Table 4, for the eight combinations together, reveals the relative extent to which the various types of thought procedure were employed by the two groups of subjects. The categories Other and Imagery (that is, other than when used in counting) are omitted -- Other, because all thought procedures reported could be classified under different categories; Imagery, because images were very infrequently employed except in counting. These two categories will also be omitted in other tables for Scottish subjects like Tables 4 and 5.

The category Indeterminate is given its own distinct place in Table 4, so designated because the nature of thought processes in some instances could not be ascertained, described, and given a more definitive name. The remaining seventeen categories are arranged under five rubrics lettered A to E, - two of them, A and E, including but a single category of thought process.

There are real advantages in classifying the categories under rubrics.

Table 4. The Number Combinations. Thought Processes with the Eight Combinations; Averages in Per Cents

Rubrics and Categories*	Cuisinaire	Conventional
Indeterminate	2	1
A. Non-Attempts	8	15
B. Counting by 1's: Objects	3	6
Images	0	1
Words Only	1	1
Partial Counting: Objects	0	1
Images	0	0
Words Only	2	3
Sub-Totals	6	12
C. Guessing	6	4
Rote Recall	10	18
Recitation of Tables	7	16
Mechanical Uses: Reverses	5	4
Inverses	7	8
Sub-Totals	35	50
D. Rational Uses: Reverses	1	1
Inverses	3	0
Inverses, Plus	2	1
Solutions, Other Forms	7	4
Sub-Totals	13	6
E. Uses of Meaningful Habituations	36	15

* Categories Other and Imagery (other than in counting) are omitted. See text.

Rubric E (Uses of Meaningful Habituations) represents the highest level of conceptual maturity and arithmetical abstraction. Rubric D comes next; but rather generally the per cents for this rubric decrease as those for Rubric E increase. At the opposite extreme of conceptual maturity are Rubrics A and B. It would be comforting and convenient to be able to say that the Rubric C categories lie at a place on the continuum intermediate between Rubrics A and B and Rubrics D and E; but one cannot be sure. Guessing, for example, may, or may not, be entirely blind: it can be governed to some extent by an awareness of number relationships, however imperfect. And tables are logical organizations, and a child who recites a table may, or may not, have some notion of its logic.

For the reasons given, major significance attaches to average per cents for Rubrics D and E as indices of advanced levels of conceptual maturity and to per cents for Rubrics A and B as indices of much lower levels. For the last-named rubrics, the larger the per cents, the less is the level of maturity. All this means that Rubric C per cents will be ignored for the most part.

To return to Table 4: the simplest way to determine the relative conceptual maturity of the thought processes used by Cui. and by Con. children is to compare the sub-totals, rubric by rubric. The most important facts to emerge are:

(a) The Con. subjects were somewhat more prone than were the Cui. subjects to refuse to attempt answers.

(b) The Con. subjects likewise made slightly more use of Counting of various kinds.

(c) The Cui. subjects used rational methods or solutions to find answers (Rubric D) somewhat more commonly than did the Con. subjects, but the average per cents for both groups are small.

(d) The average per cent of the Cui. subjects for Uses of Meaningful Habituations is more than double the size of that for the Con. subjects.* The difference in the records for Rubric E is very significant both statistically and educationally.**

(e) In Rubric C the rating of categories for level of conceptual maturity is none too certain, as has already been explained. Nevertheless, the relatively high average per cents for Rote Recall and for Recitation of Tables made by the Con. subjects raise a suspicion. This is that they had been trying to master the combinations without much understanding, and therefore at comparatively low conceptual levels.

If the suspicion mentioned in (e) could be verified as a fact, then the consistent evidence presented in (a) to (d) would be confirmed, and there would be still more reason to attribute to the Cui. program rather than to the Con. program greater effectiveness in promoting progress in conceptual maturity of procedures in ways of thinking answers for eight number combinations.

---

* Some American readers may be inclined to regard the per cents for Meaningful Habituations as small, especially for pupils at the end of their third year of schooling. If so, in their implied comparisons they will be overlooking crucial differences between American and Scottish primary education, -- differences in the age of children after three years in school, in arithmetical objectives, and the like. And, discounting these differences, they will be comparing actual data for Scottish subjects with no corresponding data at all for American children.

** Attention is directed to the phrase "significant both statistically and educationally." The requirement that pairs of per cents and of average per cents must differ by 10 points as a minimum to be regarded as significant may seem to set too high a standard. And it is too high (a) if "significant" is defined statistically to mean "reliable" and (b) if the groups compared, as in this chapter, number 200 or more subjects. Nevertheless, the standard will be held to as a guarantee that a difference noted is educationally significant; that is to say, that it identifies something of consequence for teaching and learning in the classroom.

Combinations, by operations. The gross data in Table 4 may conceal actual differences in the conceptual maturity of the groups. For example, one group might demonstrate greater maturity of thought process than the other in finding answers for the addition combinations and less maturity in finding answers for the subtraction combinations. Hence, Table 5 has been constructed to permit comparisons for pairs of combinations, operation by operation. There has been some condensation in this table when compared with Table 4, - the six kinds of Counting (three each for Counting by 1's and for Partial Counting) have been consolidated in Rubric B, and there has been some combining of categories in Rubrics C and D. None of the changes distorts differences to the advantage of either instructional group.

Addition combinations. The procedures of the Cui. subjects excel in their conceptual maturity -- about half as much Counting and double the number of answers obtained through the use of Meaningful Habituations. The much larger average per cent of answers produced by Rote Recall by the Con. subjects is noteworthy, implying (if not proving) far greater reliance upon meaningless memorization than obtained among the Cui. children.

Subtraction combinations. Inspection of the per cents for categories and sub-totals again justifies attributing superiority of conceptual maturity to the Cui. subjects, and in much the same ways as for the addition combinations.

Multiplication combinations. Once more the Cui. subjects demonstrated greater conceptual maturity, this time chiefly because of their far better record in Rubric E and partly because of their somewhat larger average per cent for Rational Solutions (Rubric D). Tentatively admissible is other evidence of the relatively less maturity of processes employed by the Con. group; namely, their greater tendency toward Rote Recall, plus their far

Table 5. The Number Combinations. Thought Processes, Operation by Operation; Averages in Per Cents

Rubrics and Categories	Operation							
	Addition		Subtraction		Multiplication		Division	
	Cui.	Con.	Cui.	Con.	Cui.	Con.	Cui.	Con.
Indeterminate	1	1	3	1	2	1	2	2
A. Non-Attempts	2	0	4	2	5	10	20	48
B. Counting, All Forms	12	22	11	25	2	1	1	0
C. Guessing	3	1	6	5	6	4	8	7
Rote Recall	16	37	11	19	6	14	7	4
Recitation of Tables	0	1	0	0	18	45	11	17
Mechanical Uses of Reverses and Inverses	2	2	18	22	10	8	16	17
Sub-Totals	21	41	35	46	40	71	42	45
D. Rational Uses of Reverses and Inverses	0	0	6	1	2	1	7	1
Solutions, Other Forms	9	8	7	8	12	5	9	1
Sub-Totals	9	8	13	9	14	6	16	2
E. Uses of Meaningful Habituations	54	27	33	18	36	11	21	4

commoner practice of reciting tables. These Rubric C categories would not be cited at all, were the data therefor not in conformity with the general picture as revealed more certainly by other data in the table.

Division combinations. One would be justified in inferring that the Con. children had had little or no instruction on these combinations. Since their average per cent for Non-Attempts is 48 and for Uses of Meaningful Habituations is 4, it is highly probable that teaching these combinations had just been started in Con. schools. Were it not for the fact that the record of the Cui. subjects is also anything but impressive, no comparison of the two groups would be made at all. For whatever the statement is worth, the Cui. subjects can be said to have attained greater conceptual maturity (small as may be the amount) in thinking answers for division combinations.

Proficiency: Accuracy and rate

Two important points made in earlier pages should perhaps be repeated. (a) The sample of eight combinations is too small to justify generalizing the findings of the Cui. and the Con subjects with respect to proficiency, to cover the 390 combinations usually taught. As here used, the term "proficiency" is restricted to the skill demonstrated in giving answers for the eight combinations in the interviews and, even more narrowly, for the pair of combinations in each operation. (b) In this study, again, measures of proficiency are of minor significance compared with measures of conceptual maturity. Proficiency data were collected and are reported merely to try to determine whether superiority in conceptual maturity means, or does not mean, a sacrifice in proficiency, -- in this section, with reference only to the eight interview number combinations.

Accuracy. The average per cents in the last column of Table 6 permit comparisons of the Cui. and the Con. subjects with respect to Accuracy with

Table 6. The Number Combinations. Proficiency in Giving Answers: Accuracy; Averages in Per Cents*

Programs and Measures	Operation				
	Addition	Subtraction	Multiplication	Division	Averages
Cuisinaire					
Correct Answers	91	77	75	65	77
Incorrect Answers	8	18	21	16	16
Non-Attempts	0	3	3	19	6
Conventional					
Correct Answers	94	77	74	34	70
Incorrect Answers	5	14	15	16	12
Non-Attempts	0	9	11	50	17

*Here and elsewhere, decimals of 0.1 to 0.5 are disregarded; decimals of 0.6 to 0.9 are rounded to 1.0

Table 7. The Number Combinations. Proficiency in Giving Answers: Rate; Averages in Per Cents*

Programs and Measures	Operation				
	Addition	Subtraction	Multiplication	Division	Averages
Cuisinaire					
Prompt Responses	74	56	57	67	63
Slight Pauses	19	24	25	16	21
Long Pauses	6	19	17	16	14
Conventional					
Prompt Responses	81	56	58	49	61
Slight Pauses	17	28	28	29	25
Long Pauses	1	15	13	22	13

*Here and elsewhere, decimals of 0.0 to 0.5 are disregarded; decimals of 0.6 to 0.9 are rounded to 1.0

the eight combinations. While the Cui. subjects have the larger average per cent for Correct Answers and the smaller average per cent for Non-Attempts, their record as a whole is not much better than that of the Con. subjects.

The two groups are equally matched in accuracy for the pairs of addition, subtraction, and multiplication combinations, but not so for the pair in division.* The substantially poorer showing of the Con. subjects with this pair of course affected adversely their average per cents for the eight combinations; for example, by decreasing their average per cent for Correct Answers and by increasing their average per cent for Non-Attempts.

Rate. The average per cents in the last column of Table 7 show no marked advantage for either group in speed of work with the eight combinations as a whole. The Con. children were able to give Prompt Responses somewhat more often to the addition combinations; the Cui. children, much more often to the division combinations. The differences between the Cui. and the Con. average per cents for Long Pauses for the pairs of combinations all amount to 6 points or less, three of them in favor of the Cui. subjects and one (for the division pair) in favor of the Con. subjects.

Summary: The number combinations

In this summary, as well as in later summaries for computation and problem solving, the results of comparisons will be brought together just as if the Cui. and the Con. subjects had been equally matched with regard to amount of instructional time, to brightness, and so on. In the chapter summary attention will be given to the actual differences in these factors and to their possible influence on the findings.

---

* The table containing the actual quantitative data has had to be omitted, as will be true for all such tables throughout the report.

In conceptual maturity the Cui. subjects clearly outmatched the Con. subjects. That is to say, they much more commonly used ways of thinking answers that are more advanced in level of meaningful abstractness, and they did so for the eight combinations as a whole and for the pairs of combinations classified by operation.

In measures of proficiency, Accuracy and Rate, there proved to be little difference between the two groups, except for minor and conflicting advantages of one group or the other with pairs of combinations.

Computation

Conceptual maturity: Ability to explain computations

However inadequate the sample of four examples may be for judging the comparative degrees of over-all computational proficiency produced by the two programs, it is regarded as large enough to warrant judgments respecting levels of ability to explain the rationale of computation. It is reasonable to assume that a child who can verbalize satisfactorily the mathematical principles involved in securing an answer for 27 + 46 could do equally well in explaining the computation in all other kinds of addition examples he has been taught to work.

Attempts to explain the rationale of the computation in each of the four examples, whether volunteered or elicited by questioning in the interviews, were classified, first, as Satisfactory or as Unsatisfactory. Within each of these major categories, explanations were further classified according to the medium used by a child: Words Only; Alt. (example altered to show a crutch, accompanied by an explanation in words); Imagery (uses of imagery; e.g., images of rods, images of the pictorial or the actual grouping of objects, etc.), and Other. In Table 8, per cents of explanations in each of the main

Table 8. Computation. Quality of Explanations of Computations, Attempts Only: Per Cents and Average Per Cents

Programs, and Quality of Explanation	Examples				Averages
	$\begin{array}{r}27\\ +46\end{array}$	$\begin{array}{r}92\\ -45\end{array}$	$\begin{array}{r}28\\ \times 3\end{array}$	$4\overline{)84}$	
Cuisinaire					
Satisfactory					
Words Only	18	15	12	31	19
Altered Examples	28	17	25	1	18
Imagery, and Other	0	1	0	1	0
Sub-Totals	46	33	37	33	37
Unsatisfactory					
Words Only	17	25	24	66	33
Altered Examples	35	40	37	1	28
Imagery, and Other	1	3	2	1	2
Sub-Totals	53	68	63	68	63
Conventional					
Satisfactory					
Words Only	3	2	3	20	7
Altered Examples	23	21	16	1	15
Imagery, and Other	0	0	0	0	0
Sub-Totals	26	23	19	21	23
Unsatisfactory					
Words Only	14	23	19	69	31
Altered Examples	60	53	61	9	46
Imagery, and Other	0	1	2	0	1
Sub-Totals	74	77	82	78	78

categories and their sub-categories, except for Imagery and Other, are summarized for Cui. and Con. subjects. The per cents for Imagery and Other are so small that they have been consolidated. No per cents could of course be entered for failures to attempt explanations.

The entries in the last column of Table 8, which are averages of the per cents for the four examples combined, may first be considered. Only 37% of all explanations attempted by Cui. subjects were classified as Satisfactory, and only 23% of those attempted by Con. subjects. As was explained in Chapter 3, to be judged Satisfactory, explanations had to be offered in terms of place value and the renaming of the numbers involved in computation; e.g., in the example, 27 + 46, 7 (ones) and 6 (ones) are 13 (ones), the 13 (ones) then to be regrouped or renamed as 1 ten and 3 ones, and so on. (However, rational explanations based upon other equally approved mathematical principles were accepted, but they were exceedingly rare.) Apparently, neither in Cui. schools nor in Con. schools had much attention been given to the understanding of computation in such terms; either that, or else experiences in explaining computations orally in this way were too few to develop skill in verbalization. Be that as it may, the record for the four examples made by the Cui. subjects materially excels that of the Con. subjects in relative frequency of Satisfactory Explanations.

In both types of school, examples were commonly altered to show crutches. The Cui. subjects altered 46% of their examples in this way; the Con. subjects, 61%. That use of crutches did not necessarily increase understanding of computations (or ability to explain computations) is shown by the fact that in only 40% of their altered examples were explanations offered by Cui. children such that they could be rated as Satisfactory; and the corresponding per cent for Con. subjects is 25.* The two groups made almost the same average per cent for Satisfactory Explanations when they altered examples; but the Cui. subjects were much the more successful in explaining computations with Words Only (15% vs 7%).

To turn to comparisons of the groups, operation by operation:** in the addition example the Cui. subjects successfully explained their computations far more commonly than did the Con. subjects, and the former subjects were especially successful in purely verbal explanations.

For the subtraction example, the per cent of the Cui. subjects for Satisfactory Explanations is 10 points larger than that of the Con. subjects.

The per cent of the Cui. H group for Satisfactory Explanations of the multiplication computation is about twice the size of that of the Con. group.

And there is a difference of 12 points between the per cents of the two groups for Satisfactory Explanations of the division computation, the difference favoring the Cui. group.

Proficiency: Accuracy and quality of performance

In computing answers for the four examples used in the study, the two groups were equally accurate (Table 9), having almost identical average per cents for Correct Answers, for Incorrect Answers, and for Non-Attempts. For the example 27 + 46 there is no advantage for either group in any of the three measures of accuracy. For 92 - 45, the Con. group has the better record as a whole. For 3 x 28, the differences in per cents are all small,

---

* The use of crutches may or may not increase proficiency in computation when the major objective is mechanical skill. When instruction stresses understanding, their role may change considerably. Crutches can enhance understanding by "picturing" what is to be done.

** The table containing the actual quantitative data has had to be omitted, as will be true for all similar tables.

Table 9. Computation. Proficiency: Accuracy and Quality of Performance: Per Cents and Average Per Cents

Programs, and Types of Measure	Examples				Averages
	27 + 46	92 - 45	28 x 3	$4/\overline{84}$	
Cuisinaire					
Accuracy					
Correct Answers	83	58	64	71	69
Incorrect Answers	16	30	27	19	23
Non-Attempts	1	12	9	10	8
Quality of Performance*					
Smooth	80	71	66	77	63
Erratic	20	29	34	23	26
Conventional					
Accuracy					
Correct Answers	84	69	60	69	70
Incorrect Answers	15	29	33	12	22
Non-Attempts	1	2	7	19	7
Quality of Performance*					
Smooth	85	74	72	83	78
Erratic	15	26	28	17	21

* Attempts only.

but consistently favor the Cui. subjects. For 84 ÷ 4, the per cents for Correct Answers are alike, while that for Incorrect Answers is larger for the Cui. subjects, and that for Non-Attempts smaller.

As for Quality of Performance, the Con. subjects excelled considerably in relative number of Smooth Performances in working the four examples (78% vs 63%), and excelled in this category in working each example, but by small margins.

Summary: Computation

The Cui. children are to be credited with having attained a higher degree of conceptual maturity than had the Con. children in the case of computation. That is to say, they were much more able than were the Con. children to demonstrate understanding of the mathematical rationale of computation by explaining the mathematical ideas and principles involved in working the four examples as a whole and in working each example.

Comparisons respecting proficiency in computation show no difference between the groups in Accuracy in working the four examples. While now one group and then the other was more accurate with this or that example, the differences are small and tend to cancel each other. In Quality of Performance the Con. subjects were consistently superior, if sometimes only slightly so. However, since Accuracy is more crucial than Quality of Performance as a criterion, the two groups are to be regarded as equally proficient in computation. To the last sentence there might be appended the cautionary phrase, "in the four interview examples."

Problem Solving

Conceptual maturity: Ability to explain solutions

The hypothesis is that, in general, children who can satisfactorily explain the reasons for their procedures in solving one-step problems are further along in conceptual maturity than are those who cannot do so. On the other hand, there is no certainty that children who cannot explain their solutions adequately thereby demonstrate low levels of conceptual maturity. Their difficulty may reside in inability to verbalize. That such may be the case sometimes is one possible inference from the fact that, in this study, so many attempted explanations of solutions had to be assigned to the category Indeterminate.*

In Table 10 the records of the Cui. and the Con. children are summarized with reference to the quality of explanations offered for solutions used with the twelve problems in the interviews. Averages in per cents permit comparisons for four categories of explanations of solutions for the twelve problems combined and for the set of problems involving each operation. The data for the category Fair will be minimized. This category was given interviewers in order to provide an "out" for them in classifying explanations that made some sense, even if little, and that,

---

* Actually, 18% of all the explanations attempted by all the Scottish children are categorized as Indeterminate, in contrast to the 1.5% of thought processes so classified in the case of the number combinations. This difference faithfully reflects the greater difficulty of evoking and interpreting the reasons for procedures adopted to solve problems, compared with ways of thinking answers for the number combinations. A very simple explanation for this greater difficulty might be the comparative inexperience of the children in talking about the process of problem solving. Probably a more basic explanation is the sophisticated nature of the ideas required to understand the meaning of each operation, this plus the complexity of the ability to identify the need for this or that operation in the language of a problem and, finally, the necessity of verbalizing the whole business in a manner intelligible to the interviewer, or to the teacher.

Table 10. Problem Solving. Reasons Given for Solutions, by Operation, Attempts Only; Averages in Per Cents

Programs, and Quality of Reasons	Problems, by Operation				
	Addition	Subtraction	Multiplication	Division	Averages
Cuisinaire					
Good	58	28	36	30	38
Fair	16	25	26	30	22
Unsatisfactory	12	24	23	22	20
Indeterminate	14	23	14	18	17
Conventional					
Good	48	30	31	21	32
Fair	23	24	32	27	26
Unsatisfactory	12	24	22	32	22
Indeterminate	17	22	16	21	19

otherwise, they would have had to record as Unsatisfactory. But the effect may well have been to increase their generosity unduly. In any case, it is probably safer to compare the groups of subjects with respect only to the widely separated categories, Good Explanations and Unsatisfactory Explanations.

The Cui. children were slightly more successful than were the Con. children in explaining solutions of the twelve problems; but the argument in their favor rests rather upon consistency than upon marked disparities between pairs of average per cents. Actually, the differences in the categories Good, Satisfactory, and Indeterminate are all so small (6, 2, and 2 points, respectively) that each would ordinarily be dismissed as insignificant and go unmentioned. However, the fact that they all point in the same direction may be reason enough to assign to the Cui. subjects a slight degree of superiority in conceptual maturity.

Before comparisons are made, operation by operation, a word concerning the problems themselves may be helpful. The three problems in each set were carefully chosen and phrased, in order (a) to insure having familiar social settings, (b) to avoid clues to specific operations, and (c) to include different kinds of problem characteristic of each given operation. The following analysis may clarify point (c) especially.

Addition

Problem 3 .... 5 chairs for visitors; then 6 more chairs for visitors .... How many visitors? (Not, How many chairs?)

Problem 6 .... 9 sums (examples) right and 7 sums wrong .... How many sums done?

Problem 12 .... 5 apples eaten, 5 apples left .... How many apples bought? ("Left," commonly a clue to subtract, but falsely here.)

Subtraction

Problem 1 .... 16 eggs in a basket; 7 cracked .... How many eggs not cracked? (Required: to find "the other part" of the total, 16.)

Problem 5 .... 14 fish in a big tank; 9 fish in a little tank .... How many more fish in the big tank? (Required: to find "the difference," or to make a comparison.)

Problem 9 .... box of 12 crayons; only 8 can be found .... How many of the crayons are missing? (Required: to find "the part gone.")

Multiplication

Problem 2 .... 5 sheets of paper to make a book .... How many sheets needed to make 5 books?

Problem 8 .... bag of marbles shared equally among 3 boys; each got 4 marbles .... How many marbles in the bag at the start? ("Shared," commonly a clue to divide, but falsely here.)

Problem 11 .... 5 boxes of eggs; 6 eggs in a box .... How many eggs all together? ("All together," commonly a clue to add; addition is appropriate here except that multiplying is more efficient.)

Division

Problem 4 .... 16 biscuits in a box; 4 boys to share them equally .... How many will each boy get? (Partition division.)

Problem 7 .... 20 children, 4 children at a table .... How many tables needed? (Measurement division.)

Problem 10 .... 18 books to be put away in equal piles of 6 .... How many piles? (Measurement division.)

However much the three addition problems differed one from another, these differences did not prevent the Cui. children from offering relatively more Good Explanations in each, by margins ranging from 9 to 13 percentage points, so that their average per cent for this category is 10 points larger than that of the Con. subjects. The average per cents of the two groups for Unsatisfactory and for Indeterminate Explanations are about the same for the set of addition problems.*

For both groups, explanations of solutions were much more difficult for subtraction than for addition problems, especially for subtraction problem 5 which requires a comparison. For this problem only one in six or seven attempted explanations could be classed as Good. In average per cents for Good Explanations the Con. children outclassed the Cui. children in two of the three subtraction problems; but in those for Unsatisfactory and Indeterminate Explanations there is no consistent advantage for either group.

For the three multiplication problems together and for each separately, the Cui. group excelled slightly in relative frequency of Good Explanations. In average per cents for Unsatisfactory and for Indeterminate Explanations for the entire set one group matches the other.

In comparative frequency of Good Explanations for the set of division problems, the Cui. subjects surpassed the Con. subjects (9 percentage points), as they did for each of the three problems, - in problem 10 (measurement division) by 17 percentage points. The Con. group has the larger average per cent for Unsatisfactory Explanations for the set as a whole and for each problem, especially for problems 7 and 10 where the differences amount to 11 and 18 points, respectively.

The results of comparing the records of the two groups in accounting for what they did in solving the sets of problems confirm the superiority of the Cui. subjects. The original statement, that the Cui. subjects were "slightly more successful" than the Con. subjects in explaining problem solutions, might justifiably be modified to read that they were "somewhat more successful," in order to indicate less doubt about the matter.

---

* The table containing the actual quantitative data has had to be omitted, as will be true for all similar tables.

Proficiency: Accuracy and performance level

In Table 11 the average per cents for Correct Answers and for levels of performance intermediate between High and Low are omitted, but they can readily be calculated from the data given.

In solutions for the twelve problems combined, the Cui. subjects seem to have been a little more accurate than were the Con. subjects. The Cui. average per cent for Correct Answers is the larger of the two by 6 points, and the Cui. average per cent for Non-Attempts is the smaller by 2 points. However, the apparent superiority of the Cui. children in Accuracy rests again upon consistency rather than upon larger differences between average per cents.

When average per cents for Correct Answers and for Non-Attempts are compared for the problems in the separate operations, one would have to say that the records of the Cui. and of the Con. children are very similar for the subtraction and multiplication sets; that the advantage of the Cui. group for the addition problems is too slight to be significant; and that in solving the division problems the Cui. subjects made much the better showing. The results of these set-by-set comparisons tend to wipe out the tentative claim of the Cui. subjects to greater accuracy, and a fairer evaluation is that the two groups were equally accurate in problem solving.

The Cui. children demonstrated some superiority in Performance Level in solving the twelve problems as a whole. Their showing in this respect is much the better of the two in solving the multiplication and the division problems; but it is just equal to the showing of the Con. subjects in solving the addition problems, and it is inferior to the latter in solving the subtraction problems.

Table 11. Problem Solving. Proficiency: Accuracy and Performance Level, by Operation; Averages in Per Cents

Programs, and Types of Measure	Problems, by Operation				
	Addition	Subtraction	Multiplication	Division	Averages
Cuisinaire					
Accuracy*					
Correct Answers	95	79	77	80	83
Non-Attempts	0	3	2	5	2
Performance Level**					
High	80	53	45	51	57
Low	5	13	18	16	13
Conventional					
Accuracy*					
Correct Answers	91	78	77	64	77
Non-Attempts	0	4	4	16	5
Performance Level**					
High	81	58	33	37	52
Low	7	18	16	29	17

* Per cents for Incorrect Answers omitted.

** Per cents for intermediate level of performance omitted.

Summary: Problem solving

The Cui. children were much more successful than were the Con. children in explaining satisfactorily solutions of addition and division problems but were no more successful than were the latter in explaining solutions of subtraction and multiplication problems. On this basis the Cui. subjects may be said to have exhibited superiority in conceptual maturity, at least to the extent indicated by saying that they were somewhat more advanced in this respect.

On the other hand, the two groups were equally proficient in solving problems when greater emphasis is placed, as it should be, on measures of Accuracy than on measures of Performance Level. After all, there is no point in performing at a higher level (as that phrase is used here) than some one else, if in the end one comes out with more incorrect answers.

Chapter Summary

Conceptual maturity

The Cui. children secured answers for number combinations by thought processes characterized by relatively high levels of meaningful abstraction much more commonly than did the Con. children.

The Cui. children were much more capable than were the Con. children of making intelligible to interviewers the mathematical rationale of computations.

The Cui. children were at least somewhat more successful than were the Con. children in explaining satisfactorily solutions of verbal problems.

The evidence collected in the investigation adds up to the conclusion that the Cui. program was considerably more effective than was the Con. program in promoting conceptual maturity. Actually, the extent of the superiority of the Cui. program in this respect is even greater than it has been shown to be in the comparisons of the findings for the two groups.

These comparisons have been made without consideration of the fact that the Cui. and the Con. groups may not have been as comparable as they have been allowed to appear. Actually, at the start of the chapter it was pointed out that there were real differences between the groups in ways which could have affected the results, -- differences which, disregarded for the time, would later receive attention. This seems to be the place to study the differences, as well as their possible influence.

According to Table 1, the Con. group, on the average, had had a total of teaching time exceeding that of the Cui. group by 25%. Moreover, the Con. subjects, according to Table 3, were estimated to be brighter than the Cui. subjects (by 0.22 of a rating) and to have succeeded better in learning arithmetic (by 0.20 of a rating) and in learning to read (by 0.24 of a rating). In a word, the Con. subjects had greater learning ability in general and had demonstrated greater ability to learn two of the school subjects, - all this, plus having had considerably more instruction.

Now, if, with these advantages, the Con. subjects had surpassed the Cui. subjects in level of conceptual maturity, there would have been little occasion for surprise. On the other hand, it would have been impossible to account for their superiority as the natural consequence of the greater effectiveness of the Con. program itself, for there would have been at hand another explanation equally as good: the Con. children were better learners and had had more time to learn.

However, the findings with regard to relative conceptual maturity favor the Cui. subjects, who had to overcome handicaps which were the counterparts of the advantages of the Con. children. The fact that they were able to do so makes their better record the more impressive.

Even so, it cannot be said that the Cui. program itself has been proved to be more effective than the Con. program. The data collected in English schools may or may not confirm the conclusion drawn from the Scottish data. If in the English schools the Cui. children should exhibit a higher degree of conceptual maturity than the Con. children, then there will be more reason to explain the difference in terms of features peculiar to the Cui. program. If, on the other hand, the Con. children in English schools should demonstrate the higher degree of conceptual maturity, then the better showing of the Cui. children in the Scottish schools cannot be ascribed to the Cui. program, *per se*. It could be, for example, that in the Scottish schools the Cui. teachers were more competent to teach the Cui. program than were the Con. teachers to teach the Con. program.

Meanwhile, there is another type of evidence, not previously considered, that seems to support the conclusion that in the Scottish schools the Cui. subjects were further along than the Con. subjects in progress toward meaningful arithmetical abstractions. This evidence lies in comparisons of the two groups with respect to their use of Altered Examples and of Words Only in explaining computations.

Both Cui. and Con. children employed crutches* extensively, except for the division example. Following is a tabulation showing, for each group of

---

* A typical crutch is shown at the right. It was of interest to see that in many schools the forms for computational practice provided, between parellel lines, spaces for entering crutches. The form which, in this study, was placed in the hands of children had no such spaces, a shortage corrected by numerous children who drew the extra line, as illustrated.

$$\begin{array}{r} 27 \\ +\,46 \\ \hline 1 \\ \hline 73 \end{array}$$

subjects and for the example in each of the three other operations, (a) the total per cent of the examples altered by children, (b) the per cent of Satisfactory Explanations with Altered Examples, and (c) the per cent (b) is of (a). (See Table 8 for the first two rows of per cents.)

	27 + 46		92 - 45		3 x 28	
	Cui.	Con.	Cui.	Con.	Cui.	Con.
(a) Per cent of examples changed to show crutches	63	83	57	74	62	77
(b) Per cent of Satisfactory Explanations in the same	28	23	17	21	25	16
(c) Per cent (b) is of (a)	44	24	30	28	40	21

Now, changing the physical appearance of an example to show a crutch provides a tangible visual clue of what is to be done. The Con. children entered crutches on their papers substantially more often than did the Cui. children, and for each of the three examples, as shown in row (a) above. One inference is that the Con. children needed the help of these visual clues more than did the Cui. children, in order to make the required computations; that is to say, the Con. subjects had been less successful than the Cui. children in moving to a more abstract kind of computation, without crutches. This inference is based on the hypothesis that crutches had been taught to the same extent in Cui. and in Con. schools, and this may not have been true. Also, it can be argued that the Con. subjects did not actually need the crutches, but used them purely as the result of a habit which had not been discouraged.

Whatever the reason, the fact remains that the Con. group was much more inclined to employ crutches than was the Cui. group. In the case of the subtraction example only (Table 9) did the Con. subjects exhibit superiority in proficiency, and their relatively greater use of crutches may have been responsible in some measure. (But if so, why not in the other two examples?) That the crutches tended, for them, to be largely mechanical devices is shown by the per cents in row (c) above.

The Cui. children as a whole altered 63% of their addition examples, and in 28% of their attempts to explain the computation in such examples they were successful. Dividing the 28% by the 63% gives the per cent 44, which is entered in row (c). This 44% is interpreted to mean that for 44% of their Altered Examples they were able to earn credit for Satisfactory Explanations. The corresponding per cent for the Con. children is 24. Comparison of the two per cents for the multiplication example, but not for the subtraction example, shows a similar advantage for the Cui. group. It appears that among the Cui. children more than among the Con. children the use of crutches fitted into (if it actually did not increase) comprehension of the rationale of computation.

Purely verbal attempts to explain computations may amount to no more than the mechanical recitation of memorized verbal formulas. Hence, the tendency of one group of children to try to explain with Words Only more commonly than does another group is no proof that the first group has demonstrated the greater degree of understanding of the mathematics of computation. When, however, one group surpasses another in per cents for Satisfactory Explanations, with Words Only, it is safe to deduce that the first-named group has attained the higher level of conceptual maturity.

Moreover, success in explaining computations reveals a greater degree of abstractness and conceptual maturity when made with Words Only than when made with the assistance of the aids furnished by crutches in Altered Examples.

The tabulation below, for Words Only, is similar to that above for Altered Examples, except that the division example can be included this time and that the wording of the categories at the left has been somewhat changed.

	27 + 46		92 - 45		3 x 28		84 ÷ 4	
	Cui.	Con.	Cui.	Con.	Cui.	Con.	Cui.	Con.
(d) Per cent of attempts to explain computations with Words Only	35	17	40	25	36	22	94	89
(e) Per cent of Satisfactory Explanations in the same	18	3	15	2	12	3	31	20
(f) Per cent (e) is of (d)	51	18	37	8	33	14	33	22

The Cui. children relied on Words Only more than did the Con. subjects in trying to explain the computation in each example, substantially so in the case of three of the examples, as shown in row (d). Far more important than the per cents in this row are those in row (f), which reveal the Cui. subjects to have had the much greater amount of success in explaining each of the four computations with Words Only.

Proficiency*

In work with each of the arithmetical tasks -- number combinations, computation, and problem solving -- the Cui. and the Con. groups were equally proficient. Two comments would seem to be in order.

(a) Granted the same amount of instructional time and equivalent amounts of ability to learn, the Cui. subjects might well have surpassed the Con. subjects in proficiency. Such re-arrangements being wholly impracticable, it is still possible to say that per unit of instructional time the Cui. children probably acquired greater proficiency than did the Con. subjects.

---

* The reader is reminded that in such statements as those of above the term "proficiency" is limited to performances with the arithmetical tasks presented in the interviews.

(b) Superiority in conceptual maturity was not purchased at the expense of proficiency; nor, on the other hand, did it increase proficiency.

It has been known for a long time that, up to the point where confusion sets in and forgetting starts, children can "learn" arithmetic when it is taught with little concern for their understanding of what they learn. It is also well known that, when meaningful abstractions are stressed, the pace of instruction must be slackened, and efforts to develop proficiency must be deferred. Exponents of the Cui. program can argue that in another year or two the Cui. children in the Scottish schools could be expected to be more proficient, and lastingly more proficient, than the Con. children. And there is no evidence to the contrary.

CHAPTER 5

FACTORS POSSIBLY INFLUENCING THE EFFECTIVENESS AND THE RELATIVE EFFECTIVENESS OF THE PROGRAMS

Several times in Chapter 4 attention was called to disparities between the Cui. and the Con. groups in averages for total time spent in learning arithmetic, for brightness, for general achievement in arithmetic, for achievement in reading, and for age. Differences in these factors are of interest in two ways.

(1) Differences in these factors, one or more of them, may have influenced the effectiveness of each of the programs, considered by itself. For example, the Cui. program may be well adapted to teach children below average in reading achievement and not be so successful with children whose achievement in this subject is above average.

(2) One program may be more effective than the other in teaching children who rate high in a given factor (e.g., brightness) and less successful than the other in teaching children who rate low in the same factor.

In this chapter in Sections 1 to 4* the influence of differences in each of four factors on each program will be examined. Comparisons on bases familiar by now will be made between H (high) and L (low) sub-groups in the larger groups of Cui. and Con. subjects, the H sub-group always having the advantage with respect to the factor under consideration, the L sub-group being at a disadvantage with respect to the same factor. Thus, in all there will be four sub-groups, or four groups as they will be referred to;

---

* There is also to be a Section 5, for age as a factor; but its organization is quite dissimilar to that for the preceding four sections.

namely, Cui. H, Cui. L, Con. H, and Con. L.

The discussion in each of Sections 1 to 4 will be organized around three questions, each of which has to do primarily with the effectiveness of the programs in encouraging children toward steadily higher levels of conceptual maturity. The data on proficiency,* though reported, are of minor importance since they will be used chiefly to ascertain whether or not progress in conceptual maturity must be purchased at the expense of progress in proficiency.

The first time around, the data pertaining to the three questions will be restricted to learning measures indicative of conceptual maturity and proficiency in each arithmetical task. Moreover, the pairs of groups furnishing the data will be viewed as if they were comparable in all significant respects, other than in the factor under study (e.g., brightness).

This plan of treatment does not of course accord with the facts. Dividing 200 and more children into two sets to constitute the H and L groups for any one of the factors did not produce as much similarity between them with respect to other factors as could have been hoped for. An examination of the contents of Table 12 will illustrate the truth of this statement for the factor, amount of instruction. Differences between the H and the L groups did exist besides those in the selected factor, time. And they occurred when H and L groups were set up for each of the three other factors.

In each of Sections 1 to 4 these differences will be disregarded for the time being, that is to say, in making the initial comparisons. In each section summary, however, they will be reported, and their possible effects on the findings will be considered.

---

* That is to say, proficiency in work with the interview items only.

Only summary tables will be presented, such as the table showing the average per cents of the H and the L groups as wholes for Correct Answers for the eight number combinations taken together. The more detailed tables, like that showing the average per cents of the two groups for Correct Answers for the pairs of number combinations, will be omitted in this chapter and in all succeeding chapters. However, data will be drawn from such unreported tables as needed and will be made use of in appropriate contexts.

Even so, Chapter 5 is long. The reader who is not particularly interested in detailed quantitative analyses may wish to know only the findings. In this case he may choose to turn directly to summaries. At the end of Section 1 there is a resumé of the results of comparisons undertaken to assess the influence of amount of instruction as a factor. And there is a similarly organized resumé of the results for each of the other three factors at the ends of Sections 2 to 4. Finally, there is a chapter summary, in which the facts obtained regarding the four factors will be brought together and organized in a different manner.

1. Amount of Instructional Time as a Factor

The children in the two Cui. schools that in the three-year span had given the largest amounts of instruction in arithmetic will constitute the Cui. H group, and those in the two Cui. schools that had given the smallest amounts of instruction will constitute the Cui. L group. In similar fashion the Con. H and the Con. L groups were set up by taking children in the two Con. schools that had allotted to arithmetic the largest, and the two that had allotted the least amounts of time. In the selected Cui. schools the H group had been taught arithmetic for 50% more time on the average than had the L group (Table 12); but in the selected Con. schools, this difference in per cent is only 10.

> A. To what extent, if at all, was the effectiveness of the Cui. program (and of the Con. program) in promoting conceptual maturity influenced by differences in time (or in amounts of instruction)?

For an answer to this question, the records of the H and L groups will be examined separately for each program. In all such intra-program comparisons, (1) will be used to designate the comparison, Cui. H vs Cui. L, and (2), to designate the comparison, Con. H vs Con. L.

(1) Cui. H vs Cui. L

Conceptual maturity. See Tables 13 to 15.

(a) Number combinations. In relative maturity of thought processes used to find answers for the eight combinations together, there is little difference between the Cui. H and the Cui. L groups, except in Uses of Meaningful Habituations. The two average per cents for Rubrics A and B combined (the rubrics rated lowest in conceptual maturity) are the same, and the average per cents for Rubric D are closely similar. But the special

Table 12. Time as a Factor. Composition
of the H and the L Groups in Each Program

Bases of Comparison	Cuisinaire		Conventional	
	H	L	H	L
Number of Subjects	60	73	123	76
Per Cents by Sex				
Male	52	51	52	54
Female	48	49	47	46
Weighted Means for				
Total Instructional Time	720	480	750	670
Age	7 yr. 8 mo.	7 yr. 8 mo.	7 yr. 11 mo.	7 yr. 7 mo.
Estimated Ratings for				
Brightness	2.60	2.79	2.16	2.52
General Achievement in Arithmetic	2.46	2.58	1.96	2.60
Achievement in Reading	2.52	2.88	2.23	2.60

significance of responses classified as belonging to Rubric E as indices of the highest level of conceptual maturity is not to be overlooked. For this rubric the average per cent of the Cui. L group is almost double that of the Cui. H group, and for this reason the Cui. L subjects may be credited with slightly greater conceptual maturity than the Cui. H subjects.

The average per cent of the Cui. L group for Uses of Meaningful Habituations is larger than that of the Cui. H group for each pair of combinations except the subtraction pair, for which the two per cents are about the same.

Table 13. The Number Combinations. Thought Processes with the Eight Combinations; Averages in Per Cents

Rubrics and Categories*	Cuisinaire		Conventional	
	H	L	H	L
A. Non-Attempts	12	8	8	27
B. Counting, All Forms	3	7	12	9
C. Guessing	7	3	5	3
Rote Recall	16	3	17	21
Recitation of Tables	0	2	18	9
Mechanical Uses of Reverses and Inverses	16	6	13	6
Sub-Totals	39	14	53	39
D. Rational Uses of				
Reverses	1	1	1	1
Inverses	2	8	0	1
Inverses, Plus	5	3	1	1
Solutions, Other Forms	14	8	2	7
Sub-Totals	22	20	4	10
E. Uses of Meaningful Habituations	22	43	21	13

* The categories Other, Imagery, and Indeterminate are omitted. The largest average per cents for Indeterminate are 4 for Cui. L and 3 for Con. L

Table 14. Computation. Quality of Explanations of Computations, Attempts Only; Averages in Per Cents*

Classification of Explanations	Cuisinaire		Conventional	
	H	L	H	L
Satisfactory				
Words Only	34	20	10	7
Altered Examples	25	11	14	28
Unsatisfactory				
Words Only	34	34	36	26
Altered Examples	6	29	49	37

* Average per cents for the category Imagery equalled 0 for all four groups. Average per cents for the category Other also equalled 0, except for the Con. L group, for which it is 6 in the case of Unsatisfactory Explanations.

(b) Computation. In 59% of their attempts the H subjects demonstrated their grasp of the mathematical rationale of their computations for the four examples combined well enough to get credit for Satisfactory Explanations. The corresponding per cent for the L subjects is 31. The H group showed greater conceptual maturity, too, whether explanations were made with Words Only or with Altered Examples.

For the four separate computations all comparisons of per cents for Satisfactory Explanations (Words Only and Altered Examples, combined) are in favor of the H group, by differences ranging from 13 points to 42 points.

(c) Problem solving. The comparisons of the Cui. H and the Cui. L groups will omit reference to average per cents for Fair Explanations, as will be done in all later comparisons. The reason has already been explained in Chapter 4; namely, the ambiguous character of this category.

Table 15. Problem Solving. Reasons Given
for Solutions, Attempts Only; Averages in Per Cents

Quality of Explanations	Cuisinaire		Conventional	
	H	L	H	L
Good	48	34	32	33
Fair	32	24	33	31
Unsatisfactory	11	25	21	20
Indeterminate	9	17	13	16

Compared with the L subjects, the H subjects attempted explanations of problem solutions which were relatively more often classified as Good and relatively less often classified as Unsatisfactory and Indeterminate. (Table 15.) The H children were therefore not only more articulate, but they also give evidence of greater understanding in selecting operations to use and in following through to answers.

The superiority in conceptual maturity manifested by the H group in explaining their solutions of the twelve problems as a whole characterizes their record for the sets of problems classified by operation. For each set their average per cent for Good Explanations is larger than that of the L subjects, and their average per cents for Unsatisfactory and Indeterminate Explanations are smaller. In all but three of these twelve comparisons the differences in average per cents amount to at least 10 points.

Proficiency. See Tables 16 to 18.

(a) Number combinations. As is shown in Table 16, the L group was more accurate than the H group in supplying answers for the eight combinations

Table 16. The Number Combinations. Proficiency in Giving Answers: Accuracy and Rate; Averages in Per Cents

Nature of the Measures	Cuisinaire		Conventional	
	H	L	H	L
Accuracy				
Correct Answers	69	79	78	58
Non-Attempts	8	9	12	15
Rate, Attempts Only				
Prompt Responses	78	58	56	55
Long Pauses	9	27	14	13

together. Moreover, the L group was consistently more accurate for the pair of combinations in each operation. In quickness of response, a relatively less critical measure of proficiency, the L group was surpassed by the H group in giving answers for the eight combinations together and for two pairs of combinations.*

(b) Computation. In solving the four examples the L subjects were more accurate than the H subjects (a significantly larger average per cent for Correct Answers and a much smaller average per cent for Non-Attempts) and surpassed them as well in Quality of Performance. The L group did no better than the H group in solving the addition and division examples, but did much better in solving the other two examples.

---

[*] In statements like this, it is to be understood that differences between per cents that amount to no more than 5 points will seldom be referred to. On the other hand, mention will always be made of differences that are significant (10 points or more) and, on occasion, of differences that are nearly significant.

Table 17. Computation. Proficiency: Accuracy and Quality of Performance; Averages in Per Cents

Types of Measure	Cuisinaire		Conventional	
	H	L	H	L
Accuracy*				
Correct Answers	54	66	71	61
Non-Attempts	21	2	5	13
Quality of Performance, Attempts Only				
Smooth	64	75	69	87
Erratic	36	25	31	13

* Average per cents for Incorrect Answers, omitted here, can easily be computed from the data given.

(c) Problem solving. The two groups were very similar in the two Accuracy categories for the twelve problems, though the H group excelled considerably in the category, High-Level Performances. (Table 18.) As for the sets of problems, the only significant differences between the average per cents of the two groups are: more Correct Answers for the H subjects in solving one set; more Non-Attempts for the H subjects in solving three sets; more instances of High-Level Performances on the part of the H subjects in all four sets. All in all, the two Cui. groups were about equally proficient in solving the one-step problems.

Table 18. Problem Solving. Proficiency: Accuracy and Performance Level; Averages in Per Cents

Types of Measure	Cuisinaire		Conventional	
	H	L	H	L
Accuracy*				
Correct Answers	82	83	80	70
Non-Attempts	4	2	3	11
Level of Performance*				
High	67	51	49	54
Low	7	17	17	20

* Average per cents for Incorrect Answers and for Indeterminate Levels of Performance, omitted here, can be readily computed from the data given.

(2) Con. H vs Con. L

Conceptual maturity. (a) The total average per cents obtained by combining the average per cents for Rubrics A and B are 20 for the H subjects and 36 for the L subjects, indicating a relatively greater tendency of the L subjects to use the lowest levels of thought processes with respect to conceptual maturity, in finding answers for the eight combinations. (Table 13.) While the L subjects used Rational Solutions slightly oftener than the H subjects, their average per cent for Rubric E is materially smaller than that of the H subjects, small as is the latter group's average per cent. On the whole, the H group exhibited a slightly higher level of conceptual maturity.

The H subjects have larger average per cents for Uses of Meaningful Habituations for all pairs of combinations, but by only 7 points in two cases.

They also counted oftener for the subtraction pair. Their average per cents for Non-Attempts are much smaller than are those of the Con. subjects for two pairs. These pair-by-pair comparisons do not increase certainty that the H group actually was superior to the L group in level of conceptual maturity of thought processes in dealing with the number combinations.

(b) The L subjects offered significantly more Satisfactory Explanations than did the H subjects when they tried to verbalize the mathematical rationale of the four computations. (Table 14.) They were much more successful in their explanations when they altered examples to show crutches.

The L subjects surpassed the H subjects in explaining satisfactorily three of the separate computations, by margins of 9 percentage points or more. However, they could do little with the division example, where a difference of 31 percentage points for Satisfactory Explanations favors the H subjects.

(c) Between the H and the L groups there is no real dissimilarity in average per cents for Good, Unsatisfactory, and Indeterminate Explanations, - this for their solutions of the twelve problems as a whole. (Table 15.)

Also, the two groups differ scarcely at all in their average per cents in these categories for the four sets of problems. For example, the average per cents for the H and L subjects in the category, Good Explanations, are within 5 points of each other; and in only one of the several comparisons for other categories are the average per cents as much as 10 percentage points apart.

Proficiency. (a) The Con. H and the Con. L groups have practically identical average per cents for Rate of response in announcing answers for the eight combinations (Table 16); but answers were much more often correct for the H subjects.

The H subjects have much the larger average per cents for Correct Answers for two pairs of combinations, and the larger average per cent for Prompt Responses for one of these pairs. On the whole, the H subjects surpassed the L subjects in proficiency with the number combinations.

(b) In solving the four examples (Table 17), the H subjects made a better record than the Con. L subjects for Accuracy (a higher average per cent for Correct Answers and a lower average per cent for Non-Attempts), but a poorer record for Quality of Performance (a smaller average per cent for Smooth Performances and a larger average per cent for Erratic Performances.

The H subjects were more accurate, too, in computing answers for three of the examples, while the L subjects had relatively more Smooth Performances for three examples. The H group is entitled to recognition for being more proficient in computation than the L group.

(c) The H group surpassed the L group in Accuracy in solving the twelve problems (Table 18), but it just equalled the latter group in Quality of Performance.

The H subjects excelled in Accuracy for the sets of multiplication and division problems. There is little difference between the two groups in Quality of Performance for any of the four sets. Accuracy being the crucial criterion of proficiency at the age of these children, in problem solving as in computation and in finding answers for the number combinations, the H subjects are judged to have been more skillful than the L subjects.

B. Was either of the programs more effective than the other in promoting conceptual maturity when the children taught in each program had had (supposedly) the same amount of instruction?

Comparisons will be made of the data, first, for the H groups in the two programs, (3) Cui. H vs Con. H; then for the L groups in the two programs, (4) Cui. L vs Con. L.

(3) Cui. H vs Con. H

Conceptual maturity. (a) In ways of thinking answers for the eight number combinations the Cui. H subjects exhibited a slightly higher level of conceptual maturity than did the Con. H subjects. This evaluation is based on relatively less Counting (a difference of 9 percentage points) and relatively greater use of Rational Solutions (a difference of 12 percentage points). The two groups employed Meaningful Habituations to the same extent. (Table 13).

The Con. subjects found many more of their answers, particularly for multiplication combinations, by reciting tables, a Rubric C category. This fact may, or may not be significant in comparisons of the conceptual maturity of the two groups, depending upon the degree of understanding present when this thought procedure was employed.

The average per cent of the Cui. group for Rational Solutions and Uses of Meaningful Habituations combined is higher than that of the Con. group for each of the four pairs of combinations, by margins ranging from 9 to 20 points.

(b) Whether they used Words Only or Altered Examples, the Cui. subjects were greatly superior to the Con. subjects in ability to explain computations for the four examples combined.

In comparisons for the separate examples the average per cent of the Cui. subjects for Satisfactory Explanations is larger in each instance to the extent of at least 20 points.

(c) In their attempts to explain their solutions of the twelve problems, the Cui. children were consistently more successful than were the Con. children. The average per cent of the Cui. subjects for Good Explanations is larger than that of the Con. subjects, and their average per cents for Unsatisfactory and Indeterminate Explanations are smaller.

For every one of the four sets of problems the average per cent for Good Explanations is larger for the Cui. group, by differences never less than 14 points. And for every set, except that in addition, the average per cent for Unsatisfactory Explanations is smaller for the Cui. group.

Proficiency. (a) Both for the eight combinations as a whole and for each pair of combinations, except the pair of division combinations, the Con. group was the more accurate of the two, though in promptness of response it was inferior. In this arithmetical task, the Con. subjects registered superior proficiency.

(b) In computing answers for the four examples together the Con. subjects demonstrated the greater Accuracy (a larger average per cent for Correct Answers and a smaller average per cent for Non-Attempts), but just equalled the Cui. subjects in Quality of Performance. The Con. group was also more accurate than the Cui. group in solving the separate examples, save for that in addition. There is no significant difference between the per cents of the two groups for Smooth Performances for any set of problems.

(c) In solving the twelve problems the two groups matched each other in Accuracy, for their average per cents for Correct Answers and for Non-Attempts are closely similar. In Level of Performance, the Cui. group has the better record. In Accuracy the two groups were equal in solving two sets of problems, but the Cui. H subjects were more accurate in solving one set and less accurate in solving another set. Comparisons of average per cents for Smooth Performances consistently favor the Cui. group.

(4) Cui. L vs Con. L

Conceptual maturity. (a) The thought processes used by the Cui. L group to obtain answers for the eight combinations rate higher in conceptual maturity than do those employed by the Con. L group. Comparisons of the average per cents of the two groups for Rubrics A, D, and E, especially those for the last-named rubric, support this statement.

For each pair of combinations the Cui. L average per cent for Rubric E is larger, never by less than 20 points. The weakness of the Con. L subjects in finding answers for the multiplication and the division combinations has been noted before. For the latter their average per cent for Uses of Meaningful Habituations is 0.

(b) The total average per cent for Satisfactory Explanations (Words Only and Altered Examples combined) is 35 for the Con. group and 31 for the Cui. group. Save for differences in such average per cents for the method preferred (Words Only and Altered Examples), the two groups seem to have been at about the same level of conceptual maturity in ability to make understandable the mathematical rationale of the four computations.

Comparisons of per cents for Satisfactory Explanations of the separate computations show no advantage for either group in one example, advantages for the Con. subjects in two examples, and a very large advantage for the Cui. group in the remaining example. The results of these comparisons strengthen the statement made above, that the groups were at about the same level of conceptual maturity in the ability in question.

(c) The Cui. L and the Con. L groups were about equally successful in explaining their solutions of the twelve problems. In only one category, Unsatisfactory Explanations, is the difference between average per cents for the two groups as large as 5 points. (Table 15.)

Likewise, in comparisons for the four sets of problems all differences between average per cents in the various categories of explanation are small, so that neither group demonstrated anything like consistent superiority over the other.

Proficiency. (a) In finding answers for the eight combinations the Cui. group outclassed the Con. group in Accuracy, having a substantially larger average per cent for Correct Answers and a slightly smaller average per cent for Non-Attempts. (Table 16.) (The difference in the last comparison is probably insignificant.) The average per cents for Prompt Responses are similar for the two groups, but that for Long Pauses is much larger for the Cui. group.

Average per cents for Correct Answers are about the same for the two groups for two pairs of combinations, but they strongly favor the Cui. subjects for the other two pairs. The advantage with respect to Prompt Responses lies with each group for two pairs of combinations.

(b) The Cui. subjects were a match for the Con. subjects in Accuracy of answers for the four examples, when both measures (Correct Answers and Non-Attempts) are taken into account; but they were much inferior in the comparisons for Quality of Performance. (Table 17.)

The per cents for Correct Answers are substantially larger for the Cui. subjects in solving two examples, but substantially smaller in solving the other two examples.

(c) the Cui. subjects were more skillful than the Con. subjects in solving the twelve problems. (Table 18.) While they just about equalled the Con. L subjects in Level of Performance, their record for Accuracy is much better. The average per cent of the Cui. group is larger, and its average per cent for Non-Attempts is smaller than the corresponding per cents for the Con. group.

The Cui. subjects have much larger average per cents for Correct Answers in solving two sets of problems, but (by barely significant margins) smaller average per cents in solving the other two sets.

C. How, if at all, was the relative effectiveness of the programs in promoting conceptual maturity influenced by differences in amount of instruction?

(5) Cui. H vs Con. L

Conceptual maturity. (a) Comparisons of the average per cents of the Cui. H and the Con. L groups for Rubrics A, D, and E reveal that the Cui. H subjects were more mature conceptually than the Con. L subjects in ways of thinking answers for the eight combinations. (Table 13.)

The average per cents of the Con. L subjects for Non-Attempts are much larger than those of the Cui. H subjects for two pairs of combinations. For all pairs except that in subtraction the average per cents of the Cui. H group for Uses of Meaningful Habituations are consistently the larger; and for all pairs, except the addition pair, their average per cents for Rational Solutions are the larger.

(b) The Cui. H group was much more successful than the Con. L group in explaining the mathematical rationale of the four computations. The total of Satisfactory Explanations represents 59% of all the attempts made by the Cui.

H subjects and 35% of all the attempts made by the Con. L subjects. (Table 14.)

In none of the comparisons for the separate examples do the Con. L subjects come as close as 18 points to the per cent for the total of Satisfactory Explanations earned by the Cui. H subjects.

(c) The Cui. H group has a substantially larger average per cent than the Con. L group for Good Explanations of their solutions of the twelve problems, as well as smaller average per cents for Unsatisfactory and Indeterminate Explanations. (Table 15.)

Similarly, the Cui. H subjects have considerably larger average per cents for Good Explanations for all four sets of problems and considerably smaller average per cents for Unsatisfactory Explanations for all sets except that in addition .

Proficiency. (a) As shown in Table 16, the Cui. H subjects were more proficient than the Con. subjects in giving answers to the eight combinations as a whole. The Cui. H group gave substantially more answers that were correct, made slightly more attempts, and responded much more quickly.

On the other hand, the Con. L group was actually somewhat more accurate than the Cui. H group, was guilty of no more Non-Attempts, and made relatively more Prompt Responses in dealing with the addition and subtraction combinations. All these bits of evidence that the Con. L subjects may have been superior in proficiency are more than wiped out by the larger amounts of proficiency shown by the Cui. H subjects with the pairs of multiplication and division combinations.

(b) The Con. L group led the Cui. H group in accuracy of computation for the four examples combined. While the differences between the average per cents of the two groups for Correct Answers and for Non-Attempts are not large, both of them are in favor of the Con. L group. The same group excelled in Quality of Performance. (Table 17.) As for Accuracy in the separate computations, the Con. L subjects were found to be somewhat inferior to the Cui. H subjects in working one example and more successful, by larger margins, in working two examples.

(c) In their attempts to solve the twelve problems, the Cui. H subjects made the better record for Accuracy (both measures). (Table 18.) The Cui. H group also excelled in Level of Performance for all four sets of problems and in Accuracy for all sets except that in subtraction.

(6) Cui. L vs Con. H

Conceptual maturity. (a) On the basis of their larger average per cents for Rational Solutions and for Uses of Meaningful Habituations (Table 13), the Cui. L subjects can be credited with using higher levels of thought processes than the Con. H subjects with the eight combinations. The far larger average per cents of the latter group for Rote Recall, for Recitation of Tables, and for Mechanical Uses of Reverses and Inverses may well point in the same direction.

For every one of the four pairs of combinations the Cui. L subjects have larger average per cents for Rational Solutions and for Uses of Meaningful Habituations. Only one of the differences is as small as 9 points.

(b) The Con. H subjects registered 24% for Satisfactory Explanations as their average for the four computations; the Cui. L subjects, 31%. (Table 14.) The differences between the per cents for Satisfactory Explanations of the separate computations are all very small, except for a 15 point advantage for the Cui. L group in the case of one example. The evidence as a whole is that neither group was superior to the other in conceptual maturity in the ability in question.

(c) For the twelve problems together, the average per cents for Good, for Unsatisfactory, and for Indeterminate Explanations are about the same for these two groups, the largest difference being only 4 points. (Table 15.) The differences in average per cents are also small in the categories Good and Unsatisfactory Explanations for the four sets of problems, one group and then the other having the advantage, but in only one set by as many as 8 points. It is fair to say that the two groups exhibited about the same degree of conceptual maturity.

Proficiency. (a) For the eight combinations the Cui. L subjects about equalled the Con. H in Rate (though charged with relatively more Long Pauses) and in average per cents for Correct Answers and for Non-Attempts. (Table 16.) The two groups have closely similar average per cents for Correct Answers for each pair of combinations, except for that in division, with which the Cui. L subjects did better.

(b) When both measures are taken into account, the Cui. L group almost, if not actually, equalled the Con. H group in Accuracy in computing answers for the four examples; and, while the differences in average per cents for Smooth and for Erratic Performances are small, both favor the Cui. L subjects. (Table 17.) The records of the two groups with respect to Accuracy and Quality of Performance show equal proficiency in computations for two of the four examples, but that of the Cui. L group is somewhat poorer for the other two examples. On the whole, neither group exhibited the greater degree of computational skill.

(c) In solving the twelve problems the Cui. L subjects made as good a showing as did the Con. H subjects with respect to Accuracy and Quality of

Performance. (Table 18.) In the set of division problems only is there a marked difference between the two average per cents for Correct Answers, and that in favor of the Cui. L group. In average per cents for High-Level Performances the Con. H group led in work with one set, and the Cui. L group, in work with two sets. Again there is no difference in the proficiency of the two groups, as there was none in the case of the other two arithmetical tasks.

Summary, Section 1

Not much success was achieved in the attempt to evaluate time as a factor in influencing the effectiveness of each program or their relative effectiveness, in encouraging children to advance to higher and higher levels of conceptual maturity. The reasons will appear.

A. How did differences in amount of instruction given to arithmetic influence the effectiveness of the Cui. program? of the Con. Program?

Comparison (1). The Cui. H group had had 50% more instructional time than the Cui. L group had had, and the H group exhibited the greater degree of conceptual maturity in their work with all three arithmetical tasks. But the H subjects were estimated to be brighter than the L subjects, by 0.19 of a rating, and to have been more successful in school learning, -- in arithmetic in general, by 0.12 of a rating, and in reading, by 0.36 of a rating. Their advantages in learning ability, by themselves, could conceivably account for the superiority of the H subjects in conceptual maturity. If so, the results of comparison (1) reveal nothing about the influence of wide differences in time upon the effectiveness of the Cui. program.

Comparison (2). The amount of instruction given the Con. H group exceeded that of the Con. L group by only 10%, and the H group outmatched the L group, and then very slightly, in but one of the arithmetical tasks. The findings in (2) therefore are of little use in providing an answer for question A. In the first place, a mere 10% difference in time is too small to have had much effect. In the second place, the L subjects were estimated to be much inferior in learning ability, having lower ratings by 0.36, 0.14, and 0.37 in brightness, arithmetic achievement, and reading achievement, respectively.

B. Was one program more effective than the other when the groups taught in each program had had the same amount of instruction?

Comparison (3). The Cui. H and the Con. H groups meet the requirement of equal amounts of instruction. The Cui. H subjects excelled slightly in level of thought processes with the number combinations and greatly in ability to explain both abstract computations and solutions of problems. The Cui. program was therefore more effective than the Con. program, the more so since the Cui. subjects had had to overcome handicaps in learning ability. Teachers estimated them to be much less bright than the Con. subjects, by 0.34 of a rating, and much less successful in learning arithmetic, by 0.50 of a rating, and in learning to read, by 0.29 of a rating.

Comparison (4). No such statement can be made concerning the relative effectiveness of the two programs when the children taught had had relatively small but equal amounts of instruction, for this requirement was not satisfied. Actually, the Con. L group had had 40% more teaching time than the Cui. L group had had.

Perhaps it should be added that, despite their large disadvantage in instructional time and despite lower ratings in brightness (by 0.27) and in reading achievement (by 0.12), the Cui. L subjects equalled the Con. L subjects in conceptual maturity in work with two of the arithmetical tasks and surpassed them considerably in work with the third.

C. How do the programs compare in effectiveness in promoting conceptual maturity when unlike amounts of instruction were given?

Comparison (5). Unfortunately, the Cui. H and the Con. L groups seem to have disregarded the clear implication of the letters H and L by having had about the same amount of instruction. Comparisons of these groups are irrelevant, therefore, as far as question C is concerned.

However, just to complete the record: the Cui. H subjects, under conditions of equal teaching time and of fairly similar ratings in other factors, surpassed the Con. H subjects in the degree of conceptual maturity they exhibited in dealing with all three of the arithmetical tasks.

Comparison (6). Compared with the Cui. L group, the Con. H group had a 55% advantage in amount of teaching time and enjoyed ratings higher by at least 0.53 in brightness and in school achievement. Yet, the Cui. L subjects equalled the Con. H subjects in ability to explain computations and solutions of problems and demonstrated the higher degree of comceptual maturity in ways of thinking answers for the number combinations. In (6), therefore, time was definitely not a factor that influenced the relative effectiveness of the two programs in the expected manner, and the least that can be said in favor of the Cui. program is that, per unit of time, it was far more successful in promoting conceptual maturity than was the Con. program.

2. Brightness as a Factor

In the absence of standard test measures of intelligence, each teacher rated his pupils for brightness on a four-point scale. Rating 1 was reserved for pupils who, in his judgment, were almost certain to be successful in the Grammar School, of whom there might be many, or few, or none in his class. Ratings 2, 3, and 4 represent decreasing amounts of estimated brightness.

To study brightness as a factor, all pupils with the rating 2 were omitted, and those with the rating 1 (the H subjects) are compared with those with ratings of 3 and 4 combined (the L subjects).

From what research has shown about the relationship between brightness and school learning, it is practically a foregone conclusion that comparisons between H and L groups in the same program, Cui. or Con., will prove to be in favor of the H groups. On the other hand, cross-comparisons (e.g., Cui. H vs Con. L) may disclose differences of interest and, possibly, of importance.

The data in Table 19 describe the composition of the four groups to be compared. Because they include both children rated 3 and children rated 4 in brightness, the L groups are larger than the H groups. The number of subjects in the Cui. H group is unfortunately quite small. Both the Con. groups had had considerably more instruction than the two Cui. groups had had. The average ages of the groups differ little. In ratings of school achievement, those for the H groups are, as could be predicted, much higher than are those of the corresponding L groups.

Table 19. Brightness as a Factor. Composition of the H and the L Groups in Each Program

Bases of Comparison	Cuisinaire		Conventional	
	H	L	H	L
Number of Subjects	45	120	67	115
Per Cents by Sex				
Male	44	49	47	53
Female	56	51	53	46
Weighted Means for				
Total Instructional Time	610	579	734	718
Age	7 yr. 9 mo.	7 yr. 9 mo.	7 yr. 10 mo.	7 yr. 10 mo.
Estimated Ratings for				
General Achievement in Arithmetic	1.22	3.20	1.19	3.03
Achievement in Reading	1.37	3.39	1.07	3.25

A. Did large differences in brightness in the children taught influence the effectiveness of the Cui. program, and of the Con. program, in promoting conceptual maturity?

(1) Cui. H vs Cui. L

Conceptual maturity. (a) As is shown in Table 20, the H subjects used higher levels of thought processes in securing answers for the eight combinations, having smaller average per cents for Non-Attempts and for Counting (not by much in the latter case), and a much larger average per cent for

Table 20. The Number Combinations. Thought Processes with the Eight Combinations; Averages in Per Cents*

Rubrics and Categories	Cuisinaire		Conventional	
	H	L	H	L
A. Non-Attempts	0	13	6	19
B. Counting, All Forms	3	9	4	17
C. Guessing	4	8	1	6
Rote Recall	7	10	18	19
Recitation of Tables	10	6	18	13
Mechanical Uses of Reverses and Inverses	15	7	17	11
Sub-Totals	36	31	54	49
D. Rational Uses of				
Reverses	0	1	1	0
Inverses	3	3	1	0
Inverses, Plus	1	3	1	1
Solutions, Other Forms	4	8	4	4
Sub-Totals	8	15	7	5
E. Uses of Meaningful Habituations	52	27	28	9

* Totals of columns do not equal 100% because of omitted categories

Rubric E. Indeed, the H average per cent for Uses of Meaningful Habituations is practically double that of the L subjects. The somewhat smaller average per cent of the H subjects for Rational Solutions may possibly be explained by

Table 21. Computation. Quality of Explanations of Computations, Attempts Only; Averages in Per Cents*

Classification of Explanations	Cuisinaire		Conventional	
	H	L	H	L
Satisfactory				
Words Only	26	13	11	3
Altered Examples	33	11	21	7
Unsatisfactory				
Words Only	19	41	23	41
Altered Examples	22	30	44	47

* Totals of columns do not equal 100% because of omitted categories of very infrequent occurrence.

their having meaningfully habituated so many of the combinations.

The H group also unmistakably, and in much the same manner, revealed the greater degree of conceptual maturity in the way they thought answers for each pair of combinations.

(b) Whether they used Words Only or Altered Examples, the H subjects were more successful than the L subjects in elucidating in the interviews, the mathematical rationale of computations in the four examples as a whole. (Table 21.) Their average per cent for Satisfactory Explanations is 59; that of the L subjects is 24.

The total per cent for Satisfactory Explanations (Words Only and Altered Examples combined) is considerably larger for the H subjects in the comparisons, example by example.

Table 22. Problem Solving. Reasons Given for Solutions, Attempts Only; Averages in Per Cents

Quality of Explanation*	Cuisinaire		Conventional	
	H	L	H	L
Good	60	27	46	22
Unsatisfactory	8	28	11	33
Indeterminate	11	19	19	17

*The category Fair is omitted because of its ambiguous character. Per cents for this category can be found from the data given.

Table 23. The Number Combinations. Proficiency in Giving Answers: Accuracy and Rate; Averages in Per Cents

Nature of the Measures	Cuisinaire		Conventional	
	H	L	H	L
Accuracy				
Correct Answers	93	68	86	65
Non-Attempts	1	10	7	19
Rate, Attempts Only				
Prompt Responses	82	51	76	53
Long Pauses	9	23	7	18

Table 24. Computation. Proficiency: Accuracy and Quality of Performance; Averages in Per Cents

Types of Measure	Cuisinaire		Conventional	
	H	L	H	L
Accuracy				
Correct Answers	88	56	91	55
Non-Attempts	3	12	0	11
Quality of Performance, Attempts Only				
Smooth	88	63	88	69
Erratic	12	37	12	31

Table 25. Problem Solving. Proficiency: Accuracy and Performance Level; Averages in Per Cents*

Types of Measure	Cuisinaire		Conventional	
	H	L	H	L
Accuracy				
Correct Answers	95	74	94	71
Non-Attempts	3	1	1	8
Level of Performance, Attempts Only				
High	70	40	73	43
Low	2	23	5	27

* None of the total per cents for Accuracy and for Level of Performance equals 100 because of omitted categories

(c) The H subjects are credited with an average per cent for Good Explanations of their solutions for the twelve problems that is more than twice the size of that for the L subjects. (Table 22.) The first-named children were also relatively more articulate, as is shown, not only by their greater success in explaining solutions, but also by their somewhat smaller average per cent for Indeterminate Explanations.

For each of the four sets of problems the average per cent for Good Solutions is at least 20 points larger for the H than for the L subjects.

Proficiency. (a) That the H group was more proficient than the L group in finding answers for the eight combinations as a whole is simply established, in Table 23, by their considerably larger average per cents for both measures of Accuracy and for both measures of Rate. For each pair of combinations the Cui. group has the larger average per cent for Correct Answers and for Prompt Responses and the smaller average per cent for Non-Attempts and for Long Pauses.

(b) In computing answers for the four examples combined (Table 24) and the answer for each example the H subjects were more proficient than the L subjects. In all comparisons of the latter kind, the margins of difference between per cents are large, on the order of the differences in the table between average per cents for the four examples together.

(c) The H group was much more skillful than the L group in solving the twelve problems. The H average per cents for Correct Answers and for High-Level Performances are considerably larger than are those of the L group; and the H average per cent for Low-Level Performances is substantially smaller. (Table 25.) In solutions of the four sets of problems, the H average per cents for Correct Answers exceed those of the L group in all four sets, in three by at least 20 points; and the H average per cents for High-Level Performances are larger in all four sets, by at least 17 points in eact set.

(2) Con. H vs Con. L

Conceptual maturity. The L group was relatively more inclined, than was the H group, to refuse to attempt answers and to use Counting to find answers for the eight number combinations. On the other hand, the H group has much the larger average per cent for Uses of Meaningful Habituations. (Table 20.)

The H subjects consistently have the larger average per cent for Rubric E in dealing with each pair of combinations, especially so in working with the division combinations. On the whole, the evidence is that the H subjects were more advanced conceptually than the L subjects in their ways of thinking answers for the number combinations.

(b) The average per cents for Satisfactory Explanations of the four computations are 32 for the H subjects and 10 for the L subjects (Table 21); and the larger per cent for Satisfactory Explanations of each of the separate computations was earned by the H subjects. None of the differences in the latter type of comparison amounts to fewer than 17 points.

(c) The H subjects surpassed the L subjects in ability adequately to explain solutions of the twelve problems. (Table 22.) The first-named subjects have much the larger average per cent for Good Explanations and much the smaller average per cent for Unsatisfactory Explanations.

The L group revealed its relative inferiority with respect to conceptual maturity by consistently having the smaller average per cent for Good Explanations and the larger average per cent for Unsatisfactory Explanations for each of the four sets of problems.

Proficiency. (a) In all four comparisons made possible by the average per cents for the eight combinations in Table 23, the H subjects have the advantage in regard to proficiency. The differences between corresponding average per cents for Correct Answers, for Non-Attempts, for Prompt Responses, and for Long Pauses are all in their favor by 11 points or more. The superior proficiency of the H subjects is revealed when comparisons are restricted to records for the four pairs of combinations classified by operation. For each pair the H group was more proficient both in Accuracy and in Rate. To cite but two types of comparison: the differences between the average per cents of the two groups for Correct Answers for the pairs of combination average 21 points; and the differences between the two pairs of average per cents for Prompt Responses average 14 points.

(b) The H subjects were much more skillful than were the L subjects in computing answers for the four examples. Their greater proficiency is shown by the large differences in their favor in each of the four pairs of average per cents summarized in Table 24. In every one of the four separate computations the L subjects made by far the poorer record in proficiency, both in Accuracy and in Quality of Performance.

(c) A cursory look at the contents of Table 24 will convince the reader that once again the H subjects surpassed the L subjects in proficiency, this time in attempts to solve the twelve problems. The differences between the average per cents for the two groups are large in amount, except in the case of Non-Attempts. All comparisons of the average per cents of the two groups for the sets of problems result in differences in favor of the H subjects in Accuracy and in Level of Performance.

B. Which program, if either, was more effective in promoting conceptual maturity when the children in both programs were, on the average, rated equal in brightness?

(3) Cui. H vs Con. H

Conceptual maturity. (a) The average per cents for Rubrics A, B, and D are about the same for the two groups. (Table 20.) However, the relatively commoner practice of the Con. H subjects in resorting to Rote Recall and to the Recitation of Tables to get answers for combinations raises the suspicion that they tended more than did the Cui. H subjects to rely on memorization rather than on understanding of relationship; but it is only a suspicion. On the other hand, the far larger average per cent of the Cui. H subjects for Uses of Meaningful Habituations warrants the belief that they were further along in the movement toward real conceptual maturity.

Evidence to support this belief, which is based on reactions to the eight combinations as a whole, is furnished by the results of comparisons for the pairs of combinations. For each pair the average per cent for Uses of Meaningful Habituations is significantly larger for the Cui. H group than for the Con. H group.

(b) With children rated 1 in brightness the Cui. program was far more successful than the Con. program in developing insight into the mathematical principles of computation, or insight plus the ability to make it intelligible to others. (Table 21.) The Cui. H subjects surpassed the Con. H subjects in average per cents for Satisfactory Explanations for the four examples combined (59 and 32, respectively) and in the per cent for such explanations for each separate example. In the latter comparisons, all differences amount to 18 points or more.

(c) The Cui. program was also more successful than the Con. program in teaching bright children how to explain solutions of problems. For the twelve problems taken together the Cui. H average per cent for Good Explanations is 14 points larger than the Con. H average per cent, and the Cui. H average per cent for Indeterminate Explanations is smaller than that of the Con. H group, this time, however, by a narrower margin. (Table 22.) And the Cui. H subjects have the larger average per cents for Good Explanations for all of the four sets of problems, by a difference of at least 10 points in each comparison.

Proficiency. (a) In finding answers for the eight combinations, the Cui. H group was more accurate and quicker in responding than the Con. H group. However, the differences between the average per cents of the two groups are rather small. (Table 23.) If the Cui. H subjects are awarded superiority in proficiency, the decision must be based rather upon the consistency than upon the size of their advantages.

But confidence in this evaluation is shaken when the records of the two groups for the pairs of combinations are compared. Then it is discovered that their special and very great difficulty with the division combinations affected adversely the average per cents of the Con. H subjects; for example, by lowering their average per cent for Correct Answers. As a consequence, the differences between Cui. H and Con. H average per cents were exaggerated. Probably the safest and fairest conclusion to draw is that the two groups demonstrated equal proficiency.

(b) According to the average per cents in Table 24, one group proved to be just as skillful as the other in computing answers for the four examples as a whole. Their average per cents for measures both of Accuracy and of Quality of Performance are remarkably similar. So are the two groups closely alike in their proficiency records for the multiplication and division examples; but the Con. H subjects have the larger per cent for Correct Answers for the subtraction example, to cancel the larger per cent of the Cui. H subjects for Correct Answers for the addition example.

(c) The two groups were also equally proficient in solving the twelve problems, as judged by the similarity of their average per cents for Accuracy and for Quality of Performance in Table 25. They were equally proficient, too, in solving the sets of problems. In these last-named comparisons, both for Accuracy and for Quality of Performance, only one difference between the pairs of average per cents is as large as 8 points, and the rest amount to 5 points or fewer.

(4) Cui. L vs Con. L

Conceptual maturity. (a) With children rated 3 and 4 in brightness, the Cui. program was more successful than the Con. program in encouraging the discovery and adoption of steadily higher conceptual levels of ways of thinking answers for the number combinations. The Cui. L group is charged with relatively fewer Non-Attempts with the eight combinations and with less use of Counting. The differences between the average per cents of the groups for these categories are not large; but if they are combined, the total average per cents for Rubric A and Rubric B are 22 for the Cui. L subjects and 36 for the Con. L subjects. And this combining of categories is justified on the ground that they represent the most immature ways of dealing with number combinations. Though both groups made limited use of Meaningful Habituations, the average per cent of the Cui. H group is much the larger of the two. (Table 20.)

For each pair of combinations the Cui. L subjects surpassed the Con. L subjects in average per cent of responses classed as belonging to Rubric E. The purport of all these comparisons is to confirm the truth of the first sentence in the foregoing paragraph.

(b) The children in neither of these groups could explain many of their computations satisfactorily; but those taught according to the Cui. program were able to do so more often than were those taught according to the Con. program. For the four examples combined, an average of 24% of the Cui. L attempts were classified as Satisfactory Explanations, compared with an average of 10% of the Con. L attempts. (Table 21.)

For each separate example the Cui. L per cent for Satisfactory Explanations proved to be at least double the size of the Con. L per cent.

(c) The Cui. L subjects explained solutions of the twelve problems little, if any, better than did the Con. L subjects. (Table 22.)

For only two sets of problems are the Cui. L average per cents for Good Explanations significantly larger than those of the Con. L subjects. The pairs of average per cents for Unsatisfactory Explanations are about the same in size for all sets except that in division, with which the average per cent of the Con. L group is much the larger.

Proficiency. (a) One group was just about as proficient as the other in finding answers for the eight combinations. In only one of the four comparisons that can be made with the average per cents in Table 23 is there much difference, namely, in Non-Attempts. Otherwise, that is, in relative frequencies of Correct Answers, Prompt Responses, and Long Pauses, the two groups are about alike.

Of the sixteen comparisons for pairs of combinations (four pairs, four measures) only in five are differences between pairs of average per cents as large as 9 points, three of them in favor of the Con. L subjects. In the process of averaging, however, their far poorer record with the division combinations wiped out their advantages elsewhere.

(b) Neither group excelled in average per cents for Correct Answers and for Non-Attempts in solving the four examples; but the Con. L subjects computed steadily and smoothly slightly more commonly. (Table 24.)

The per cents for the two groups in the categories Correct Answers, Non-Attempts and Smooth Performances differ only once by as many as 10 points in comparisons, example by example; and the differences are credited to neither group consistently enough to change the judgment that the Cui. L and the Con. L groups were equally proficient in computation.

(c) On the basis of their average per cents for Accuracy and for Level of Performance as shown in Table 25, one would have to say that these two groups were alike in ability to solve the twelve problems. In solving the multiplication set the Con. L subjects were considerably more successful, but only in terms of the one criterion, Correct Answers. Contrariwise, the Cui. L subjects were more successful in solving the division set, this time in terms of three categories: Correct Answers, Non-Attempts, and High-Level Performances.

C. When the children taught according to one program were brighter than the children taught according to the other program, how did the Cui. and the Con. programs compare with respect to effectiveness in promoting conceptual maturity?

(5) Cui. H vs Con. L

Conceptual maturity. (a) Comparisons of the average per cents of these two groups with respect to Rubrics A, B, and E in Table 20 establish the higher conceptual level of the ways in which the Cui. H subjects thought answers for the eight combinations.

There are some comparisons for the pairs of combinations in which the average per cents of the Con. L group equal those of the Cui. H group, but never in use of Meaningful Habituations. For Rubric E the Cui. average per cents regularly are at least five times as large as are those of the Con. L group.

(b) The Cui. H subjects were overwhelmingly superior to the Con. L subjects in ability to explain computations, having almost five times as many attempts, on the average, judged to be Satisfactory Explanations for the four computations combined. (Table 21.) The Cui. H per cent for Satisfactory Explanations for each separate computation is regularly at least three times the corresponding Con. L per cent.

(c) The Cui. H subjects greatly surpassed the Con. L subjects in successful attempts to explain solutions, both for the twelve problems as a whole (Table 22) and for each set of problems. In comparisons of the latter type the difference between the two average per cents for Good Explanations is always 30 points or more.

Proficiency. (a) For the eight number combinations as a whole, the results of comparisons of average per cents for Accuracy and for Rate all favor the Cui. H group (Table 23), who also excelled in the same way for each of the four pairs of combinations.

(b) In working the four examples the Cui. H group was more competent than the Con. L group, both in Accuracy and in Quality of Performance (Table 24.) For each of the separate examples, the same group has the larger per cents for Correct Answers, by margins of 20 points or more, as well as for Smooth Performances, by margins of at least 15 points.

(c) The Cui. H subjects were much more competent in solving the twelve problems as a whole and in solving each set of problems. (Table 25.) Comparisons of average per cents give decisive advantages to the Cui. H group for all four sets in the category Correct Answers, for two sets in the category Non-Attempts, and for one set in the category High-Level Performances. Other differences between pairs of average per cents are non-existent or too small to be reported.

(6) Cui. L vs Con. H

Conceptual maturity. (a) The Cui. L group held up surprisingly well in comparison with the Con. H group with respect to level of thought processes employed with the number combinations. (Table 20.) True, the Cui. L average per cent for Rubrics A and B together is 22 and that of the Con. H group is 10, but the Cui. L subjects relied relatively less often on Rote Recall and the *Recitation of Tables for answers, excelled somewhat in the use of Rational* Solutions, and equalled the Con. H subjects in average per cent for Uses of Meaningful Habituations. On the whole, the two groups seem to have been at about the same stage of conceptual maturity.

This judgment is confirmed by the results of comparisons for pairs of combinations. Neither group consistently has the larger average per cent for Uses of Meaningful Habituations. On the contrary, each group has the larger average per cent for two pairs.

(b) Of the Cui. L attempts to explain the four computations, an average of 24% were acceptably good, compared with an average of 32% of the Con. H attempts. (Table 21.) This difference is not significant; but all differences between per cents of the two groups for Satisfactory Explanations of the separate computations are to be credited to the Con. H subjects, three of them

large enough to be significant. The results of these last comparisons strengthen the case of the Con. H group for having the higher level of conceptual maturity in this instance.

(c) The average per cent of the Con. H group for Good Explanations is larger, and that for Unsatisfactory Explanations is smaller than the corresponding average per cent of the Cui. L subjects. (Table 22.) The Con. H subjects were without exception more successful in explaining solutions for the sets of problems, the difference in average per cents for Good Explanations always being 12 points or more.

Proficiency. (a) The Con. H group surpassed the Cui. L group in proficiency in giving answers for the eight combinations. The Con. H average per cents for Correct Answers and for Prompt Responses are considerably larger than those for the Cui. L group; and the Con. H average per cents for Long Pauses and for Non-Attempts are smaller, but in the last comparison by a narrower margin. (Table 23.) For all pairs of combinations the Con. H subjects have larger average per cents for Correct Answers and also for Prompt Responses for all pairs except that in division.

(b) All four comparisons of the average per cents summarized in Table 24 show the Con. H subjects to have been much more skillful than the Cui. L subjects in computing answers for the four examples. The Con. H per cents for the separate examples are all much larger for Correct Answers and for Smooth Performances, and they are smaller for Non-Attempts in all except the division example.

(c) The Con. H group demonstrated the greater degree of skill in solving the twelve problems, a fact established by comparing the average per cents of the two groups for Correct Answers and for High-Level and Low-Level Performances in Table 25. The Con. H subjects were more proficient as well in solving the sets of problems, -- all four sets in terms of relative frequency of Correct Answers and of High-Level Performances.

Summary, Section 2

A. It makes sense to suppose that a program of instruction will be more effective in promoting conceptual maturity when the children taught rate high rather than relatively low in brightness. The expected results were obtained both for the Cui. and for the Con. program.

In both (1) and (2) the groups compared had had about the same amount of instruction, but in both the H groups had much higher ratings than the L groups in general achievement in arithmetic and in reading achievement. While these latter differences in favor of the H groups may have affected the results, they are best interpreted as the usual concomitants of their superior brightness, and their influence has not been separated from that of brightness.

B. It is not so easy to predict which of two programs will be more effective in promoting conceptual maturity when the children taught according to both programs rate either high in brightness or relatively low in brightness. The data examined in Section 2 show, in (3), that the Cui. program was the more effective of the two in teaching children rated 1 in brightness, and was so in all three arithmetical tasks. The data in (4) show that the Cui. program was more effective than the Con. program in teaching children rated 3 and 4 in brightness, but this time only in encouraging higher levels of processes in thinking answers for the number combinations and in developing greater ability to explain computations in terms of their mathematical rationale. It was no more effective, however, in teaching children how to make their solutions of verbal problems intelligible to others.

In comparison (3) the Cui. subjects had had 17% less teaching time in arithmetic than the Con. subjects, and in comparison (4), the Cui. subjects had had 19% less instruction. Moreover, in both comparisons the Con. subjects enjoyed the higher ratings in over-all achievement in arithmetic and in reading achievement, substantially so in three of the four comparisons. In view of these facts the Cui. program is seen to have an even stronger claim to greater effectiveness than was shown by the raw comparisons first reported.

C. Probably one can foretell with some assurance that a given program of instruction, when taught to children rated high in brightness, will be more effective in promoting conceptual maturity than another program, when taught to children rated 3 and 4 in brightness. And in this study the predictions would be quite accurate, for in (5) the Cui. H subjects clearly surpassed the Con. L subjects in level of conceptual maturity in all three arithmetical tasks, and in (6) the Con. H subjects surpassed the Cui. L subjects in two of the three.

In (5) the Cui. H subjects, with their much higher ratings in brightness and in school learning, apparently easily surmounted a disadvantage of 15% in instructional time; and in (6) the Cui. L subjects, along with their much lower ratings in brightness and school learning, had to contend with 20% less instruction than the Con. subjects had had. In (6), then, the Cui. L subjects did well to equal the Con. H level of conceptual maturity in even one arithmetical task.

3. General Arithmetic Achievement as a Factor

All of the children rated 1 in over-all achievement in arithmetic in the Cui. schools make up the Cui. H group; and all those rated 3 and 4, the Cui. L group. In the same way the Con. H and the Con. L groups were constituted from all the Con. schools.

As when H and L groups were chosen with respect to time and to brightness as factors, the four groups selected with respect to over-all achievement in arithmetic are unequal in the amount of instruction they had received. (Table 26.)

The two H groups have ratings close to 1.00 in brightness and ratings actually equal to 1.00 in reading achievement; and the two L groups have ratings of 3.26 and 3.11 in brightness and of 3.43 and 3.38 in reading achievement. So closely related are the three factors -- brightness, reading achievement, and arithmetic achievement -- that to group children according to one factor is to group them almost as accurately according to each of the other two factors. The results of comparisons in this section should therefore resemble those in Section 2.

For another reason, the findings of the comparisons to be reported shortly may be predictable in large measure. After all, children were rated for achievement in arithmetic and they were "tested" in arithmetic. And while the teachers who furnished ratings for their pupils did so with no knowledge of their relative success in the interviews, their ratings must have been influenced by their classroom observations of their pupils' relative success in reacting to the very arithmetical tasks included in this investigation.

Table 26. General Achievement in Arithmetic as a Factor.
Composition of the H and the L Groups in Each Program

Bases of Comparison	Cuisinaire		Conventional	
	H	L	H	L
Number of Subjects	42	129	74	124
Per Cents by Sex				
Male	55	45	49	51
Female	44	56	53	46
Weighted Means for				
Total Instructional Time	627	577	733	718
Age	7 yr. 9 mo.	7 yr. 9 mo.	7 yr. 10 mo.	7 yr. 10 mo.
Estimated Ratings for				
Brightness	1.30	3.26	1.17	3.11
Achievement in Reading	1.00	3.43	1.00	3.38

A. Did differences in over-all arithmetic achievement influence the effectiveness of the Cui. program, and of the Con. program, in promoting conceptual maturity?

(1) Cui. H vs Cui. L

Conceptual maturity. (a) The H group employed ways of thinking answers for the eight combinations that are substantially more advanced conceptually than are those used by the L group. (Table 27.) The H subjects attempted 100% of the answers; the L subjects, 87%; and the H average per cent for Rubric E

Table 27. The Number Combinations. Thought Processes with the Eight Combinations; Averages in Per Cents

Rubrics and Categories	Cuisinaire		Conventional	
	H	L	H	L
A. Non-Attempts	0	13	6	18
B. Counting, All Forms	2	9	5	16
C. Guessing	4	8	2	6
Rote Recall	10	11	16	19
Recitation of Tables	12	6	17	14
Mechanical Uses of Reverses and Inverses	15	9	19	9
Sub-Totals	41	34	54	48
D. Rational Uses of				
Reverses	0	0	1	0
Inverses	4	3	1	0
Solutions, Other Forms*	5	12	5	5
Sub-Totals	9	15	7	5
E. Uses of Meaningful Habituations	47	27	26	9

* The category, Inverses, Plus, is included

is 20 points larger than that of the L subjects. The H average per cent for Uses of Meaningful Habituations is consistently the larger of the two for each pair of combinations, insignificantly so for the multiplication pair but by margins of 18 points or more for the other three pairs.

Table 28. Computation. Quality of Explanations of Computations, Attempts Only; Averages in Per Cents*

Classification of Explanations	Cuisinaire		Conventional	
	H	L	H	L
Satisfactory				
Words Only	32	13	12	3
Altered Examples	31	14	19	9
Unsatisfactory				
Words Only	15	41	24	39
Altered Examples	22	31	45	46

* Not all totals of columns equal 100% because of omitted categories of very infrequent occurrence.

Table 29. Problem Solving. Reasons Given for Solutions, Attempts Only; Averages in Per Cents

Quality of Explanation*	Cuisinaire		Conventional	
	H	L	H	L
Good	60	27	48	23
Unsatisfactory	10	27	12	31
Indeterminate	13	19	19	16

* The category, Fair, is omitted.

Table 30. The Number Combinations. Proficiency in Giving Answers: Accuracy and Rate; Averages in Per Cents

Nature of the Measures	Cuisinaire		Conventional	
	H	L	H	L
Accuracy				
Correct Answers	93	68	86	64
Non-Attempts	0	10	8	18
Rate, Attempts Only				
Prompt Responses	85	52	76	54
Long Pauses	4	21	7	18

(b) The average per cents for Satisfactory Explanations of the four computations are 63 for the H group and 27 for the L group. (Table 28.) For only one of the four examples, the subtraction example, does the L per cent for Satisfactory Explanations come as close as 31 points to the corresponding H per cent.

(c) According to Table 29, the average per cent of the H group for Good Explanations is double the size of the corresponding average per cent of the L group, and the H average per cent for Unsatisfactory Explanations is almost a third that of the L group. For each set of problems the H average per cent for Satisfactory Explanations is at least 30 points larger than that of the L group.

Proficiency. (a) The H group excelled considerably in all comparisons for Accuracy and for Rate in finding answers for the eight combinations (Table 30); and, for the pairs of combinations, all average per cents for Correct Answers are larger by 13 to 39 points for the H subjects, and all average per cents for

Table 31. Computation. Proficiency: Accuracy and Quality of Performance; Averages in Per Cents

Types of Measure	Cuisinaire		Conventional	
	H	L	H	L
Accuracy				
Correct Answers	90	57	91	56
Non-Attempts	4	11	0	11
Quality of Performance, Attempts Only				
Smooth	92	63	86	70
Erratic	8	37	13	29

Table 32. Problem Solving. Proficiency: Accuracy and Performance Level; Averages in Per Cents

Types of Measure	Cuisinaire		Conventional	
	H	L	H	L
Accuracy				
Correct Answers	95	74	94	67
Non-Attempts	0	5	1	7
Level of Performance, Attempts Only				
High	77	44	74	45
Low	4	20	5	27

Prompt Responses are larger for the H subjects by 30 points or more.

(b) In solving the four examples (Table 31), the H subjects on the average attempted more computations, had far more Correct Answers, and worked much more smoothly and steadily than did the L subjects. For the separate examples the per cents of the H subjects for Correct Answers exceed those of the L subjects by margins never less than 20 points, and their per cents for Prompt Responses, by margins never less than 19 points.

(c) Comparisons of the average per cents for the H and the L groups in Table 32 show that the H group was much more skillful in solving the twelve problems as a whole. The average per cent of the H subjects for Correct Answers is larger than that of the L subjects by 9 points only for the addition problems, but their average per cents for Correct Answers are at least 25 points larger for each of the other three sets. The H average per cent for High-Level Performances is 11 points larger than that of the L subjects in solving the addition problems, but their average per cents for this category are at least 25 points larger in solving each of the other three sets.

(2) Con. H vs Con. L

Conceptual maturity. (a) As all comparisons of average per cents for Rubrics A, B, and E reveal (Table 27), the H subjects tended to think answers for the eight combinations at conceptual levels higher than those characteristic of the L subjects. The latter subjects refused to attempt answers and counted more commonly than the H subjects, significantly so for two pairs of combinations. While neither group employed Meaningful Habituations very often, all differences in average per cents for Rubric E favor the H group, by only 7 points for the division pair but by 13 points or more for each of the other three pairs.

(b) Neither group was adept in accounting mathematically for the computations of the four examples; but the average per cent of the H subjects for Satisfactory Explanations is the larger of the two, - 31 compared with 12 (Table 28). So is the H per cent for Satisfactory Explanations the larger of the two for each separate computation, - by 11 points or more for each example.

(c) Problem solving. The H subjects excelled in ability to offer Good Explanations for solutions of problems, both for the twelve problems as a whole (Table 29) and for all four sets of problems. The differences between the average per cents of the two groups for Good Explanations amount to 20 points or more for each set of problems.

Proficiency. (a) to (c). The H group has far the better record of the two for Accuracy and for Rate of work with the eight combinations (Table 30), for Accuracy and for Quality of Performance in solving the four examples (Table 31), and for Accuracy and Level of Performance in solving the twelve problems (Table 32). Pair by pair of combinations, the H average per cents exceed the L average per cents for Correct Answers, by 18 points or more for three pairs, and for Prompt Responses, by 13 points or more for three pairs. Example by example, the H per cents are the larger of the two for Correct Answers, by 23 points or more in each instance, and for Smooth Performances, by 23 points or more in each example. Set by set of problems, the H average per cents for Correct Answers are the larger, by 13 points or more for each set, and for High-Level Performances, by 11 points or more for each set.

B. How did the two programs compare with respect to effectiveness in promoting conceptual maturity when the children taught rated high in over-all arithmetic achievement? when they rated low?

(3) Cui. H vs Con. H

Conceptual maturity. (a) The average per cent of the Con. H subjects for Non-Attempts and Counting combined is somewhat larger than that of the Cui. H subjects, - 11 compared with 2; and the Con. H subjects used Rubric C thought processes considerably more often. (Table 27.) If the Rubric C categories do represent intermediate levels of conceptual maturity, -- and there is no certainty that they do, -- then to the extent of a 13-point difference in average per cents relatively more of the Con. H thought processes were at the half-way mark. The difference in the comparative frequencies of the two groups in Uses of Meaningful Habituations is large, and its implication is quite unambiguous. On the evidence as a whole, superiority in conceptual level of ways of thinking answers for the eight combinations belongs to the Cui. H group.

The average per cents of the two groups for Uses of Meaningful Habituations with the pairs of combinations are as follows. The first per cent in each pair is that of the Cui. H subjects.

Addition	80;	42	Subtraction	45;	32
Multiplication	34;	21	Division	27;	9

(b) As the average per cents in Table 28 show, the Cui. H group was much more able than was the Con. H group to explain the mathematical rationale of the four computations as a whole; and, as other data show, the Cui. H per cent for Satisfactory Explanations exceeds that of the Con. H subjects for each separate computation by 25 points or more.

(c) The Cui. program was more successful than was the Con. program in teaching children rated high in over-all achievement in arithmetic to make intelligible to others their methods of solving problems. Both for the twelve problems as a whole (Table 29) and for each set of problems the Cui. H subjects are credited with relatively more Good Explanations. The differences in average per cents for this category are larger for the Cui. H group for each set, and significantly so for three sets.

Proficiency. (a) If both measures of Accuracy are taken into account, the Cui. record in this respect is slightly the better of the two, as far as the eight combinations as a group are concerned. The Cui. H margin of superiority in Rate of response is somewhat larger. However, differences between average per cents of the two groups for Correct Answers are small for three pairs of combinations, and for Prompt Responses are small for two pairs. The Con. H subjects did especially poorly with the division pair, with consequent adverse effects upon their averages as reported in Table 30. On this account, the Cui. H and the Con. H groups had best be rated as equal in proficiency with the number combinations.

(b) In skill in working the four examples as a whole the two groups were about a match for each other. The near-equality of the pairs of average per cents in Table 31 accurately reflects the similarity of the records of the two groups in solving each of the four examples. For only one example, addition, is the difference between the Cui. H and the Con. H average per cents for Correct Answers as large as 10 points, the others amounting to 3 points or less; and for only one example do their average per cents for Smooth Performances differ by as many as 10 points; for the others, by 6 points or less.

(c) Likewise, there is nothing to choose between the degrees of proficiency revealed by the two groups in solving the twelve problems, with respect either to Accuracy or to Level of Performance (Table 32). The difference between the two average per cents for Correct Answers is as large as 7 points for only one set of problems; and for High-Level Performances the differences are only as large as 5 points for two sets.

(4) Cui. L vs Con. L

Conceptual maturity. (a) For the eight combinations the average per cents of the two groups for Non-Attempts and Counting combined are 22 for the Cui. L subjects and 34 for the Con. L subjects. The Cui. L average per cents both for Rational Solutions and for Uses of Meaningful Habituations are three times the size of such average per cents for the Con. L subjects. (Table 27.)

The average per cents for Non-Attempts and for Counting are significantly the smaller for the Cui. L subjects for two pairs of combinations; and the Cui. L subjects have the larger average per cent for Rational Solutions for three pairs, and for Uses of Meaningful Habituations for all four pairs.

(b) The Cui. L group was much more able than the Con. L group to account mathematically for what they did in computing answers for the four examples (Table 28) as well as for each separate example. Differences between the per cents for Satisfactory Explanations, all in favor of the Cui. L subjects, range from 25 points for the subtraction example to 38 points for the addition example.

(c) There is no advantage for either group with respect to success in explaining solutions of the twelve problems taken together (Table 29) or in explaining solutions for each of the four sets of problems. All differences between the average per cents of the two groups for Good Explanations, set by set, amount to 4 points or less.

Proficiency. (a) Despite a somewhat stronger tendency on the part of the Con. L subjects to refuse to attempt answers, the proficiency records of the two groups are about on a par for Accuracy and for Rate of work with the eight number combinations. (Table 30.) Comparisons of the average per cents of the two groups for Correct Answers for the pairs of combinations show only one significant difference (26 points), and that in favor of the Cui. L group; and this large difference cancelled the advantages, small but consistent, of the Con. L group for the three other pairs. In comparisons of Prompt Responses the large advantage of the Con. L subjects with the addition pair is cancelled by the large advantage of the Cui. L subjects with the division pair.

(b) The Cui. L and the Con. L groups were equally proficient in computing answers for the four examples as far as Accuracy is concerned, but the latter group excelled slightly in Quality of Performance. (Table 31.) The per cents of the two groups for Correct Answers differ as much as 7 points for only one example, and those for Prompt Responses are significantly larger for the Con. L subjects for two examples.

(c) On the whole, the records of the Cui. L and the Con. L groups are about the same with respect to Accuracy and Level of Performance in solving the twelve problems (Table 32); and of the eight comparisons of relative frequencies of Correct Answers and of High-Level Performances for the four sets of problems only one shows a significant difference.

C. How did the programs compare for effectiveness in promoting conceptual maturity when the children taught according to one program rated high and the children taught according to the other program rated relatively low in over-all achievement in arithmetic?

(5) Cui. H vs Con. L

Conceptual maturity. (a) The Cui. H subjects used thought processes to obtain answers for the eight combinations that are far more advanced conceptually than are those of the Con. L subjects. This fact is practically self-evident in Table 27, and it is supported by the results of comparisons for the pairs of combinations, for three of which the Cui. H subjects made significantly greater use of Meaningful Habituations.

(b) The Cui. H subjects exhibited considerably more ability to explain the computations for the four examples combined (Table 28) and for all of the separate examples. In the latter type of comparison the differences in relative frequency of Satisfactory Explanations are very large.

(c) Both for the twelve problems as a whole (Table 29) and for each set of problems the Cui. H subjects have much the larger average per cents for Good Explanations. The difference between the two average per cents amounts to 28 points or more for each of the four sets of problems.

Proficiency. (a) As shown in Table 30, the Cui. H group substantially surpassed the Con. L group in Accuracy and in Rate of response for the eight combinations as a whole. The Cui. H subjects have an advantage of 19 points or more in average per cent for Correct Answers and of 35 points or more in average per cent for Prompt Responses, in each case for three pairs of combinations. The exception is the addition pair in both instances.

(b) In solving the four examples combined, the Cui. H group demonstrated far greater Accuracy and a much better record with respect to Quality of Performance (Table 31). For each of the four separate computations the Cui. H average per cent for Correct Answers is the larger by 25 points or more, and for High-Level Performances, by 14 points or more.

(c) The Cui. H subjects were much more skillful than the Con. L subjects in solving the twelve problems (Table 32) and in solving each set of problems. The Cui. H average per cent for Correct Answers is the larger of the two by 10 points for one set of problems and by 25 points or more for the other three sets. And the Cui. H average per cent for High-Level Performances is the larger by 11 points for one set and by 25 points or more for the other sets.

(6) Cui. L vs Con. H

Conceptual maturity. (a) The two groups revealed similar rather than dissimilar degrees of conceptual maturity in the ways they thought answers for the eight combinations. For the rubric, Uses of Meaningful Habituations, the two average per cents are practically identical (Table 27), largely because the Con. H subjects' comparatively larger average per cent for this rubric for the pair of multiplication combinations is offset by the Cui. L subjects' larger per cent for the subtraction pair.

(b) There proved to be little difference between the groups in ability to supply Satisfactory Explanations for the four computations taken together (Table 28) and for the four separate computations. Only for the division example is there a significant difference between the two per cents for Satisfactory Explanations. This difference is in favor of the Con. H subjects.

(c) The Con. H subjects excelled considerably in explaining satisfactorily solutions for the twelve problems (Table 29), and their average per cents for three sets of problems are larger than those of the Cui. L subjects by 19 points or more.

Proficiency. (a) In supplying answers for the eight number combinations the Con. H subjects excelled by large margins both in Accuracy and in Rate of response (Table 30); and their average per cents for Correct Answers are 12 or more points the larger for three pairs of combinations, and for Prompt Responses their average per cents are 30 or more points the larger for three pairs.

(b) The Con. H subjects were more accurate than the Cui. L subjects in working the four examples and much more commonly did so smoothly and steadily (Table 31). The eight comparisons of per cents for Correct Answers and for Smooth Performances all indicate the superior proficiency of the Con. H subjects in working the separate examples. No difference is smaller than 18 points.

(c) Greater proficiency in solving the twelve problems was shown by the Con. H group (Table 32), which also has the larger average per cent for Correct Answers for three sets of problems (by 19 points or more), and for High-Level Performances for all four sets (by 13 points or more).

Summary, Section 3

A. According to the results of comparisons (1) and (2), marked differences in over-all achievement in arithmetic would seem to have influenced the effectiveness of both programs in promoting conceptual maturity. Compared with the L group in each program, its paired H group exhibited much the higher levels of mental processes in thinking answers for number combinations and much more ability to explain mathematically computations in abstract examples and to make clear to interviewers methods of solving one-step problems.

The substantial differences between the H and the L groups in conceptual maturity are in no sense explicable in terms of disparities in amounts of

instruction, for the H group had the benefit of only 8% more teaching time in (1) and of only 2% more in (2). Nor, on the other hand, are the differences in conceptual maturity attributable exclusively to the substantial superiority of the H groups in over-all achievement in arithmetic, as might be supposed.

As has been reported in Table 26, the teacher ratings of the H and L groups in factors other than over-all achievement in arithmetic are as follows.

	Cuisinaire		**Conventional**	
Brightness	**H, 1.30**	**L, 3.26**	**H, 1.17**	**L, 3.11**
Achievement in Reading	**H, 1.00**	**L, 3.43**	**H, 1.00**	**L, 3.38**

On the 4-point scale used, the differences in the ratings of the H and L groups in each program are very large indeed, so large as to offer a better explanation of the differences in conceptual maturity than do the differences in arithmetic achievement in general. On this ground, comparisons (1) and (2) throw no light on the question of the influence of the last-named factor on the effectiveness of either of the two programs, when considered by itself.

B. Comparison (3) was intended to determine the relative effectiveness of the two programs when used with children rated 1.00 in arithmetic achievement; comparison (4), when the two programs were used with children rated 3 and 4.

In (3) the Cui. program was found to be the more successful in promoting conceptual maturity, for the Cui. H subjects demonstrated superiority in this respect in work with all three arithmetical tasks. To do so, they had to overcome the effects of having had 15% less instruction than the Con. H subjects had had and of being somewhat less bright than the latter, by 0.13 of a rating. In view of these facts there can be little question about the greater effectiveness of the Gui. program when taught to good achievers in arithmetic.

In (4) the Cui. program again proved to be more effective than the Con. program, but in two instead of all three of the arithmetical tasks. Yet, the Con. L subjects had had 24% more instruction than the Cui. L subjects and were judged to be somewhat brighter (0.15 of a rating). Because of these disadvantages on the part of the Cui. L group, the Cui. program is to be regarded as having been much more successful than the Con. program in promoting conceptual maturity when both programs were taught to children rated 3 and 4 in arithmetic achievement in general.

C. Comparisons (5) and (6) involve Cui. and Con. groups which in each instance have markedly different ratings in over-all achievement in arithmetic.

The greater effectiveness of the Cui. program in comparison (5) -- Cui. H vs Con. L -- could have been anticipated even though the Con. L subjects had had 14% more teaching time. However, the record of the Cui. program in producing the greater degree of conceptual maturity in work with all three arithmetical tasks probably resulted less from the fact that the Cui. subjects were much the better achievers in arithmetic in general than from the fact that they were also much brighter than the Con. subjects and were much more skilled in reading. (The differences in ratings in these factors are 1.81 in brightness and 2.38 in reading achievement).

In comparison (6) -- Cui. L vs Con. H -- the Con. program might have been expected to have been generally the more effective of the two programs, especially since the Con. H subjects had a 27% advantage in amount of instruction and enjoyed other substantial advantages, not merely in estimated over-all achievement in arithmetic, but in brightness and in reading ability (differences of 2.09 and of 2.43, respectively, in ratings). Actually, however, instead of promoting the higher degree of conceptual maturity in work with all three arithmetical tasks, the Con. program did so only for one. Under the conditions just described, the success of the Cui. program is the more impressive.

From what has been written under C., it is obvious that in neither (5) nor (6) was over-all achievement in arithmetic the principal factor in accounting for the results of comparisons. Whatever its influence, it was clouded by differences in other factors whose effects may well have been decisive.

4. Reading Achievement as a Factor

The children rated 1 in reading achievement in the Cui. and in the Con. schools make up the Cui. H and the Con. H groups, respectively; and the children rated 3 and 4 in the Cui. schools and in the Con. schools, the Cui. L and the Con. L groups, respectively. Data concerning the four groups to be compared are assembled in Table 33.

Table 33. Achievement in Reading as a Factor.
Composition of the H and the L Groups in Each Program

Bases of Comparison	Cuisinaire		Conventional	
	H	L	H	L
Number of Subjects	48	105	76	102
Per Cents by Sex				
Male	48	49	46	57
Female	52	41	54	43
Weighted Means for				
Total Instructional Time	611	583	737	713
Age	7 yr. 9 mo.	7 yr. 10 mo.	7 yr. 10 mo.	7 yr. 10 mo.
Estimated Ratings for				
Brightness	1.25	3.38	1.34	3.17
Over-all Achievement in Arithmetic	1.00	3.48	1.00	3.39

As was true in Section 3 where the H and the L groups were established on the basis of over-all achievement in arithmetic, so here where the basis is reading achievement, the H groups have high ratings, 1.25 and 1.34, and the L groups have much lower ratings in brightness, 3.38 and 3.17. And just as the H subjects according to arithmetic achievement were on the average excellent readers and the L subjects, much poorer readers, so the H subjects according to reading achievement have the highest possible rating, and the L subjects have ratings of 3.48 and 3.39 in arithmetic achievement.

Because of the closeness of the relationship between brightness, reading achievement, and arithmetic achievement, the results of the comparisons in this section should be very similar to those in Section 3.

A. How, if at all, was the effectiveness of each program influenced by marked differences in reading achievement?

(1) Cui. H vs Cui. L

Conceptual maturity. (a) The L group exhibited much the lower degree of conceptual maturity in level of thought processes used to find answers both for the eight combinations as a whole (Table 34) and for the pairs of combinations. To cite but a single category, Uses of Meaningful Habituations, the H subjects have the larger average per cent for all pairs, significantly so for three.

(b) The H subjects were far more successful than were the L subjects in accounting mathematically for the computations in the four examples combined (Table 35) and for the computation for each of the examples. The difference between the two per cents for Satisfactory Explanations, each in favor of the H group, amounts always to 20 points or more.

Table 34. The Number Combinations. Thought Processes with the Eight Combinations; Averages in Per Cents

Rubrics and Categories	Cuisinaire		Conventional	
	H	L	H	L
A. Non-Attempts	1	13	8	18
B. Counting, All Forms	2	9	7	16
C. Guessing	2	9	2	6
Rote Recall	10	11	18	20
Recitation of Tables	10	6	18	15
Mechanical Uses of Reverses and Inverses	16	8	17	10
Sub-Totals	38	34	55	51
D. Rational Uses of				
Reverses	0	1	1	0
Inverses	2	7	1	0
Solutions, Other Forms	6	9	5	6
Sub-Totals	8	17	7	6
E. Uses of Meaningful Habituations	50	23	23	8

(c) Likewise, the H subjects displayed more ability to explain solutions of problems than did the L subjects. The H average per cent for Good Explanations for the twelve solutions is more than double that of the L subjects (Table 36), and the average per cent of the H subjects for this category exceeds

Table 35. Computation. Quality of Explanations of Computations, Attempts Only; Averages in Per Cents

Classification of Explanations	Cuisinaire		Conventional	
	H	L	H	L
Satisfactory				
Words Only	25	14	13	3
Altered Examples	31	11	20	10
Unsatisfactory				
Words Only	22	41	25	37
Altered Examples	21	29	42	47

Table 36. Problem Solving. Reasons Given for Solutions, Attempts Only; Averages in Per Cents

Quality of Explanation	Cuisinaire		Conventional	
	H	L	H	L
Good	60	26	45	22
Unsatisfactory	7	30	13	32
Indeterminate	11	19	18	19

that of the L subjects by 10 points or more for each of the four sets of problems.

Table 37. The Number Combinations. Proficiency in Giving Answers: Accuracy and Rate, Averages in Per Cents

Nature of the Measures	Cuisinaire		Conventional	
	H	L	H	L
Accuracy				
Correct Answers	93	67	83	66
Non-Attempts	2	10	9	17
Rate, Attempts Only				
Prompt Responses	81	49	73	54
Long Pauses	6	23	7	18

Table 38. Computation. Proficiency: Accuracy and Quality of Performance; Averages in Per Cents

Types of Measure	Cuisinaire		Conventional	
	H	L	H	L
Accuracy				
Correct Answers	89	54	88	58
Non-Attempts	4	13	1	12
Quality of Performance, Attempts Only				
Smooth	93	61	86	70
Erratic	7	39	13	30

Table 39. Problem Solving. Proficiency: Accuracy and Performance Level; Averages in Per Cents

Types of Measure	Cuisinaire		Conventional	
	H	L	H	L
Accuracy				
Correct Answers	96	72	91	69
Non-Attempts	0	5	1	7
Level of Performance, Attempts Only				
High	71	44	70	48
Low	2	26	10	26

Proficiency. (a) The H subjects were much more accurate in their answers for the eight combinations than were the L subjects, and gave their answers more promptly. (Table 37.) And the same statement may be made for each pair of combinations. For each pair, the H average per cent for Correct Answers is the larger by at least 18 points, and for Prompt Responses, by even larger margins.

(b) The data in Table 38 show that the H group has far the better record for Accuracy and for Quality of Performance in computing answers for the four examples. Also, all differences between per cents for Correct Answers favor the H group by 20 points or more for each example, and for Smooth Performances by larger amounts.

(c) The H subjects greatly surpassed the L subjects in Accuracy and in Performance Level in solving the twelve problems as a whole. (Table 30.)

The H average per cents for Correct Answers and for High-Level Performances are larger than those of the L group for each set of problems, in only one of the eight comparisons by fewer than 10 points.

(2) Con. H vs Con. L

Conceptual maturity. (a) The L group is charged with significantly more Non-Attempts, counted (almost significantly) more commonly, and made much less use of Meaningful Habituations to get answers for the eight combinations. (Table 34.) In ten of the twelve comparisons for Non-Attempts, Counting, and Uses of Meaningful Habituations for the four pairs of combinations, the H subjects show to advantage in terms of conceptual maturity, significantly so in six.

(b) In ability to explain the mathematics of the four computations, the H group was greatly superior to the L group (Table 35), as it was also in explaining the separate computations. For each example, the H per cent for Satisfactory Explanations is at least twice the size of that for the L group.

(c) The H subjects were much the more successful in explaining solutions of problems, - of the twelve problems as a whole (Table 36) and of each set of problems. In the latter case no difference between average per cents for Good Explanations amounts to fewer than 20 points.

Proficiency. (a) Both in Accuracy of answers and in Rate of response the H subjects were much superior to the L subjects in reacting to the eight combinations together (Table 37) and in reacting to each pair. In comparisons respecting Correct Answers and Prompt Responses, pair by pair, the differences, all favoring the H group, are substantial.

(b) Compared with the L group, the H group obtained considerably more Correct Answers and worked steadily and smoothly more often in completing the

four computations. (Table 38.) And the H group excelled greatly both in Accuracy and in Quality of Performance for each separate computation.

(c) In problem solving the H subjects were clearly more proficient than the L subjects, having the better record in Accuracy and in Performance Level for the twelve problems (Table 39) and for each set of problems. All eight comparisons of the latter type show the H subjects to have been superior in proficiency, and by wide margins in seven.

B. How did the two programs compare with respect to effectiveness in promoting conceptual maturity when the children taught were excellent readers? when they were relatively poor readers?

(3) Cui. H vs Con. H

Conceptual maturity. (a) The Cui. H subjects displayed the higher degree of conceptual maturity in the ways they thought answers for the number combinations. For the eight combinations as a whole, the Con. H subjects' average per cent for Non-Attempts and for Counting combined is 15, and that of the Cui. H subjects is 3; and the Con. H average per cent for Uses of Meaningful Habituations is less than half as large as that of the Cui. H subjects. (Table 34.) For each pair of combinations the Cui. H subjects made use of Meaningful Habituations substantially more often than did the Con. H subjects.

(b) The Cui. H subjects have by far the better record in explaining computations. For the four examples their average per cent for Satisfactory Explanations, 56, exceeds that of the Con. H subjects by 23 points (Table 35), and their per cents for this category are 14 points or more the larger for each of three examples.

(c) Much greater ability to explain solutions of problems was shown by the Cui. H group, for the twelve problems together (Table 36) and for each set of problems. The difference between the average per cents for Good Explanations

for each set is significant and is in favor of the Cui. H subjects.

Proficiency. (a) When only the averages for the eight combinations are compared (Table 37), the Cui. H subjects seem to have been more proficient in work with the number combinations than were the Con. H subjects. However, for reasons to be reported below, the two groups had best be regarded as equal in proficiency.

(b) and (c) The two groups were also equally skillful in computing answers for the four examples (Table 38), certainly as far as Accuracy is concerned, and in solving the twelve problems. (Table 39.)

When the per cents or average per cents of the two groups were compared in greater detail, significant differences appeared as follows:

Number combinations; Correct Answers, 1 pair; Prompt Responses, 1 pair. These two differences, occurring in the division pair, are wholly responsible for the larger average per cents credited to the Cui. H group.

Computations: Correct Answers, 4 examples, 3 of them for the Cui. H group; Smooth Performances, no example.

Problem solving: Correct Answers, 1 set; High-Level Performances, no set.

(4) Cui. L vs Con. L

Conceptual maturity. (a) The Cui. L group has somewhat smaller average per cents for Non-Attempts and for Counting than has the Con. L group, and a substantially larger average per cent for Uses of Meaningful Habituations, these comparisons being for the eight combinations taken together. (Table 34.) For each pair of combinations the Cui. L average per cent for the last-named category is well over double the size of the Con. L average per cent, and for all pairs the Con. L group counted relatively more frequently, -- for two pairs by substantial margins.

(b) The average per cents for Satisfactory Explanations of the four computations are 13 for the Con. L subjects and 25 for the Cui. L subjects (Table 35), who also have a significantly larger per cent for Satisfactory Explanations for each separate computation.

(c) Neither of the groups revealed the higher degree of ability to explain solutions of the twelve problems (Table 36) or solutions of the problems in three sets. In the fourth set (addition) the Cui. L subjects were much more successful than the Con. L subjects.

Proficiency. (a), (b), (c) The two groups were equally accurate and prompt in responding to the eight combinations (Table 37), and they were equally skillful in working the four examples (though the Con. L subjects are credited with relatively more Smooth Performances) (Table 38) and in solving the twelve problems (Table 39).

Following are the only significant differences between the average per cents or per cents for the two groups.

Number combinations: Correct Answers, 1 pair; Prompt Responses, 1 pair.

Computations: Correct Answers, 1 example; Smooth Performances, 2 examples.

Problem solving: Correct Answers, 1 set; High-Level Performances, 2 sets.

C. How did the programs compare with respect to effectiveness in promoting conceptual maturity when the children taught according to one program were better readers than those taught according to the other program?

(5) Cui. H vs Con. L

Conceptual maturity. (a) The Cui. H subjects displayed much the higher degree of conceptual maturity in ways of thinking answers for the eight combinations as a whole (Table 34) and for the pairs of combinations. Their average per cent for Uses of Meaningful Habituations is at least four times that of the Con. L subjects for each pair.

(b) The average per cent of the Cui. H subjects for Satisfactory Explanations of the four computations is 56, compared with 13 for the Con. L subjects (Table 35), and the Cui. H per cent for this category is 15 points or more the larger for each separate computation.

(c) The Cui. H subjects greatly surpassed the Con. L subjects in ability to explain solutions of the twelve problems (Table 36) and solutions of each of the four sets of problems, - in the latter case, by differences amounting to 30 points or more between the pairs of average per cents for Good Explanations.

Proficiency. (a), (b), (c) Both in Accuracy and in Rate (or Quality or Level of Performance) the Cui. H revealed far the greater degree of proficiency in each of the three arithmetical tasks. (See Tables 37 - 39.)

The per cents or average per cents of the two groups differ significantly -- and in favor of the Cui. H group -- as follows:

Number combinations: Correct Answers, 3 pairs; Prompt Responses, 3 pairs.

Computations: Correct Answers, all 4 examples; Smooth Performances, all 4 examples.

Problem solving: Correct Answers, 4 sets; High-Level Performances, 3 sets.

(6) Cui. L vs Con. H

Conceptual maturity. (a) The thought processes of the two groups in finding answers for the eight combinations were, on the average, at about the same level of conceptual maturity. While the average per cent of the

Con. H group for Non-Attempts and Counting combined is 7 points smaller than that of the Cui. L group, its average per cent for Rational Solutions is 10 points smaller. (Table 34.) Both groups made use of Meaningful Habituations to the same extent.

Comparisons of average per cents for the four pairs of combinations show no significant differences for Non-Attempts or for Counting for a single pair, and significant differences for Rubrics D and E combined for two pairs, both in favor of the Cui. L subjects.

(b) The Con. H subjects were somewhat superior in ability to explain the four computations (Table 35), and they have the larger average per cent for Satisfactory Explanations for each of the separate computations, in only two by significant differences.

(c) Much the better record in explaining solutions of problems was made by the Con. H subjects, whose average per cents for Good Explanations are significantly larger than those of the Cui. L subjects for the twelve problems as a whole (Table 36) and for each of the four sets of problems.

Proficiency. (a), (b), (c) The Con. H group showed far greater proficiency than did the Cui. L group in dealing with each arithmetical task, -- with the eight combinations, both in Accuracy and in Rate; with the four examples, both in Accuracy and in Quality of Performance; and with the twelve problems, both in Accuracy and in Performance Level. (See Tables 37 to 39.)

Summary, Section 4

A. Comparisons (1) and (2). Each program was more effective in promoting conceptual maturity when the children taught were excellent readers than when they were poor readers. Since the H subjects in both programs had very great advantages over the L subjects in brightness and in arithmetic achievement (Table 33), it is improbable that differences in reading achievement had much to do with determining the results of the comparisons.

B. Comparisons (3) and (4). The Cui. program was more effective than the Con. program both when the subjects in the two programs were excellent readers and when they were below-average readers. The only exception occurred in (4) in the case of problem solving.

The superior records made by the Cui. subjects in both (3) and (4) are especially noteworthy in view of their having had, in each instance, about 17% less instruction than the Con. subjects. Otherwise, the groups were fairly comparable. In (3) the two H groups had ratings in brightness differing by only 0.9 of a rating and in arithmetic achievement differing not at all. In (4) the Con. L group was rated 0.21 lower in brightness than the Cui. L group, but 0.09 higher in arithmetic achievement.

C. Comparisons (5) and (6). In (5) the Cui. program was more effective than the Con. program in developing conceptual maturity in work with all three arithmetical tasks, when it was taught to excellent readers and when the Con. program was taught to poor readers. In (6) the Cui. program was the less effective of the programs when it was taught to poor readers and the Con. program was taught to excellent readers.

In (5) the Con. L group had had 22% more teaching time than the Cui. H group had had, and it may have profited from this advantage, but certainly not enough to make up for their relative shortages in reading achievement, in brightness (1.92 on the rating scale), and in arithmetic achievement (2.39). In (6) the Con. H subjects, besides being better readers than the Cui. L

subjects, had had 21% more instruction, were also brighter (2.04) and better achievers in arithmetic (2.48).

The explanation of the results in (5) and (6) is not to be found in differences in amounts of arithmetic instruction, however; nor, as a matter of fact, is it to be found in differences in reading ability. Rather, it is to be found in the fact that the group that excelled in conceptual skill was brighter than the other group and abler in school learning.

5. Age as a Factor

Large differences in the age of children taught might be expected to influence in some way the effectiveness of each of the two programs, as well as their relative effectiveness. It was hoped that the data collected in the Scottish schools might shed light on this problem; but they do not.

It will be recalled that children were clasified by age according to a four-fold scheme, starting with the bracket 7 years 5 months and less and progressing by six-month intervals to 8 years 6 months and over. Now, if the children had distributed themselves fairly evenly among the four intervals, it would have been practicable to section them into H and L groups, taking, for the L groups, those who were less than 7 years 6 months old, and for the H groups, those who were 8 years 6 months old, or older.

However, the Scottish children did not distribute themselves so conveniently. Instead, according to Table 1, about 90% of the total sample of subjects fall in the two intermediate age categories. In the case of the Cui. schools, only 4% (nine children) are in the lowest age category and only 1% (2 children) are in the oldest age category. The corresponding per cents for the Con. schools are 1 (three children) and 13 (thirty-two children). It is obvious that no fruitful information could be obtained by comparing groups numbering so few subjects.

To have combined the numbers of subjects in the two lowest age categories to constitute the L groups and those in the two highest age categories to constitute the H groups would have yielded H and L groups differing on the average little more than six months, - a difference too small to have served a useful purpose. An average difference of at least a full year between H and L groups would probably be required in order to make profitable any attempt to assess the influence of age differences on the effectiveness and the relative effectiveness of the Cui. and the Con. programs.

Because of the method used to classify children according to age, no procedure could be discovered to get H and L age groups differing in age to the extent of a year. So, regrettably, the relation of differences in age to the effectiveness of the programs must be left for study in other research specifically designed to investigate the problem.

Chapter Summary

An attempt has been made in this chapter to assess the influence of certain factors on the effectiveness of the Cui. and of the Con. programs and on their relative effectiveness; that is to say, effectiveness in advancing children to steadily higher levels of meaningful arithmetical abstractions. The factors examined were: 1. Time (total amount of instruction in the three-year period); 2. Brightness; 3. Over-all achievement in arithmetic; 4. Reading achievement, and 5. Age of subjects. Measures for the factors 2 to 4 were obtained by using teachers' ratings on a four-point scale.

It did not prove feasible, however, to evaluate age as a factor because of the impossibility of establishing suitable groups for the needed comparisons. The influence of each of the other four factors has been studied in accordance with the same plan.

(1) Within the Cui. schools and within the Con. schools "high," or H groups, and "low," or L groups were set up separately for each factor; e.g., brightness. Four sub-groups, or groups, were thus made available, designated as Cui. H, Cui. L, Con. H, and Con. L, from which six pairs of groups could be arranged for comparisons: (1) Cui. H vs Cui. L; (2) Con. H

vs Con. L; (3) Cui. H vs Con. H; (4) Cui. L vs Con. L; (5) Cui. H vs Con. L, and (6) Cui. L vs Con. H.

(2) The data for each factor were then analyzed to secure answers for three questions.

> A, To what extent, if at all, was the effectiveness of each program in promoting conceptual maturity influenced by differences (in the given factor)?

The results of comparisons (1) and (2) as described above are relevant to this question as it may be raised for each of the four factors.

> B. Was either of the programs more effective than the other in promoting conceptual maturity when the children taught in both programs were equal with respect to (the given factor)?

Here comparisons (3) and (4) should provide pertinent data.

> C. How, if at all, was the relative effectiveness of the two programs in promoting conceptual maturity influenced by differences in the children taught with respect to (the given factor)?

The results of comparisons (5) and (6) should be useful in answering this question.

The findings for each of the four factors will be treated together as they relate to the three questions, starting with question A. It should be stated at once that definitive answers to the questions were not often obtainable, for reasons that will appear.

After this summary for the three questions has been presented, other data not previously mentioned but relevant to the general problem of maturity of arithmetical concepts will be reviewed.

Questions A, B, and C

Question A. To what extent, if at all, was the effectiveness of each program in promoting conceptual maturity influenced by differences (in each factor)?

Time. It had been supposed that the Cui. H and the Cui. L subjects, and the Con. H and the Con. L subjects, would have received quite dissimilar amounts of instruction in arithmetic. This expectation was not realized in (2), for the Con. H and the Con. L groups differed in amount of instruction, not by the 20% or so desirable for meaningful comparisons, but by only 10%. On this account the results of comparison (2), while useful for other purpose, can contribute nothing toward an answer to question A.

Nor, as a matter of fact, do the results of comparison (1), but for another reason. The Cui. H subjects had had 50% more instruction than the Cui. L subjects -- and this was good -- and they demonstrated much the higher level of conceptual maturity in dealing with two of the arithmetical tasks and a slightly higher level with the third (the number combinations). On the other hand, they enjoyed such large advantages in other factors -- in brightness and in proven ability in school learning -- that these, without their extra instruction, could have been responsible for their better showing. Hence, the data reviewed provide no answer for question B, in so far as the factor, time, is concerned.

Brightness. In comparisons (1) and (2), the H group in each program had arrived at a higher level of conceptual maturity than the corresponding L group, and in all three arithmetical tasks. Both H groups had had 2% to 8% more teaching time than their paired L groups, but the differences, considered over the three-year period, are too small to have been influential in determining

the results.

But, accompanying their much higher ratings in brightness, are higher ratings for the H groups in over-all achievement in arithmetic and in reading achievement, as could have been predicted. Since, short of sophisticated statistical analysis (something the writer is loathe to undertake), the influence of brightness per se cannot be separated from the influence of greater success in school learning, about all that one can say is that each program was more effective when used with children rated 1 in brightness than when used with children rated 3 and 4.

Over-all achievement in arithmetic. The Cui. H children exhibited in comparison (1) a much higher status in conceptual maturity than did the Cui. L children, and in each of the three arithmetical tasks. While their greater success in learning arithmetic in general, apart from everything else, could possibly account for the difference, it must be recognized that the Cui. H subjects were regarded by teachers as much brighter and as much better readers than the Cui. L subjects.

In comparison (2) the Con. H group surpassed the Con. L group in the level of conceptual maturity exhibited in all three arithmetical tasks. But was the Con. H superiority in this respect the consequence of their being generally much more successful in learning arithmetic, - this, and this only? It is doubtful that this is true, for the Con. H subjects probably benefited greatly from their being much brighter and, possibly, from their being much better readers than were the Con. L subjects.

The evidence for the greater effectiveness of the two programs when the children taught rated high rather than low in over-all achievement in arithmetic is not convincing. As noted, differences in other factors complicate the problem, and the relative influence of these factors has not been determined, so that no confident answer can be given to question A. The probability is that the superiority of the H groups in conceptual maturity is attributable less to their being better achievers in arithmetic than to their being better learners in general.

Reading achievement. No great significance can be attached to comparisons of groups constituted with reference to reading achievement, either here for (1) and (2), or later for (3) to (6). The reason is the close relationship between reading achievement on the one hand and brightness and general achievement in arithmetic on the other hand. As shown in Table 33, the Cui. subjects rated 1 in reading achievement were rated 1.25 in brightness and 1.00 in over-all arithmetic achievement; and the Cui. subjects rated 3 and 4 in reading achievement were rated 3.38 in brightness and 3.48 in over-all achievement in arithmetic. The situation in the case of the Con. H and the Con. L subjects is very similar.

Nevertheless, for whatever they are worth, and largely because the information is available, it can be reported that the H groups greatly excelled in conceptual maturity in all three arithmetical tasks. But, in compliance with the need for an answer to question A, it cannot be said that reading achievement, by itself, was really a factor in determining the effectiveness of either of the programs in promoting conceptual maturity.

Question B. Was either of the programs more effective than the other in promoting conceptual maturity when the children taught according to the two programs were equal with respect to (the given factor)?

Time. The Con. L subjects having had 40% more teaching time than the Cui. L subjects, the requirement of equal amounts of instruction for the two groups to be compared cannot be satisfied. Hence, the results of comparison (4), Cui. L vs Con. L, cannot be used in connection with question B. They will contribute, however, at another point.

In comparison (3), however, the Cui. H and the Con. H groups had had approximately the same amount of instruction. The Cui. H subjects made a slightly better record for conceptual maturity of thought processes used with the number combinations than did the Con. H subjects, and they made far better records than the Con. H subjects in ability to explain computations and solutions of problems. In view of the fact that the Con. H children were rated higher in brightness and in academic achievement, the Cui. program must be credited with the greater effectiveness at the point in question.

Brightness. The Cui. program was far more successful than the Con. program in stimulating and guiding children rated 1 in brightness to move toward higher levels of conceptual maturity. Judged by the quality of their performances in the interviews, the Cui. H subjects were conceptually much more advanced than the Con. H subjects in all three arithmetical tasks, -- comparison (3).

The Cui. program was also more successful than the Con. program in forwarding progress toward conceptual maturity when the children taught rated 3 and 4 in brightness. In comparison (4) the Cui. L subjects were superior in comparisons for but two of the tasks, -- the number combinations and explanations of computations.

The fact that the Cui. H subjects were handicapped by having had 17% less teaching time than the Con. H subjects, and the Cui. L subjects, by having 19% less teaching time than the Con. L subjects, strengthens the claim of the Cui. program for greater effectiveness in promoting conceptual maturity, both for children rated 1 and for children rated 3 and 4 in brightness.

Over-all achievement in arithmetic. As shown in comparison (3), children rated high in over-all achievement in arithmetic revealed superior levels of conceptual maturity in all three arithmetical tasks when taught according to the Cui. program than when taught according to the Con. program.

Children rated relatively low in over-all achievement in arithmetic were more mature conceptually in two of the arithmetical tasks when taught according to the Cui. program. The exception occurred in ability to explain solutions of problems. See (4).

The greater effectiveness of the Cui. program in teaching children rated 1 in arithmetic achievement as well as in teaching children rated 3 and 4 is made the more impressive when one realizes that in (3) the Cui. H subjects had had 14% less instruction than the Con. H subjects and were rated as somewhat poorer readers, and that in (4) the Cui. L subjects had had 20% less instruction than the Con. L subjects and were rated somewhat lower in brightness.

Reading achievement. In (3) the Cui. H group surpassed the Con. H group, which had had 20% more instruction, in level of conceptual maturity displayed in work with the three arithmetical tasks. In (4), the Cui. L group, again faced by a relative shortage in instruction, this time amounting to 22%, demonstrated greater conceptual maturity than the Con. L group in work with two of the tasks.

Since, save for the differences in teaching time, the groups were fairly comparable, the Cui. program in both (3) and (4) proved to be more effective than the Con. program; that is to say, the Cui. program was the more effective both when the two programs were used with excellent readers and when they were used with poor readers.

Question C. How, if at all, was the relative effectiveness of the programs in promoting conceptual maturity influenced by differences in (the given factor)?

Time. Again one of the two comparisons is lost, for in (5) the Cui. H and the Con. L groups differed only 5% to 7% in the total time they had been taught arithmetic.

In comparison (6), the data show that the Cui. L group tended to think answers for the number combinations in ways more advanced conceptually than did the Con. H group; but the two groups exhibited about the same level of conceptual maturity in explaining computations and solutions of problems. Since the Con. H subjects profited at least to some extent from having had 56% more instruction than the Cui. L group and since they were brighter and superior in school learning as judged by teachers, the Cui. program should probably be rated as more effective than the Con. program per unit of instructional time.

Brightness. According to the results of comparison (5), the Cui. program, when taught to bright children, was more effective than the Con. program, when taught to less bright children, in promoting conceptual maturity in each of the three arithmetical tasks. Any other finding would be contrary to expectation, especially since the groups compared differed so much, not only in brightness, but in other factors as well. The 20% advantage of the Con. L subjects in relative amount of instruction was not enough to enable them to overcome their shortages in other respects.

In (6), the Con. program, when taught to bright children, was more effective than the Cui. program, when taught to less bright children, but this time in only two arithmetical tasks. In addition to their great advantage in brightness, the Con. H subjects had been given 27% more teaching time in arithmetic than the

Cui. L subjects, and they of course surpassed the Cui. L subjects in relative standing with regard to other factors.

In full recognition of the influence of other factors, it is still possible to say with some confidence that brightness was definitely a factor in determining the effectiveness of each of the programs.

Over-all achievement in arithmetic. According to the results of comparison (5), the Cui. H group, despite having had 13% less instruction, exceeded the Con. L group in level of conceptual maturity displayed in all three arithmetical tasks.

The findings of comparison (6) reveal that the Con. H subjects, with an advantage of 27% in amount of instruction, were more mature conceptually than the Cui. L subjects in only one of the arithmetical tasks; namely, in ability to explain solutions of problems.

The findings in (5) are what would have been anticipated from the higher ratings of the Cui. H subjects, not only in over-all arithmetic achievement, but in brightness and in reading achievement. Even their relative deficiency in time would be unlikely to have lessened materially the effect of their advantages in other factors.

The findings in (6), on the other hand, might not have been anticipated in view of the fact that the Con. H subjects, besides having had much more instruction than the Cui. L subjects, also outranked them in other factors, especially brightness. The Cui. L children did exceptionally well to equal the Con. H subjects in two arithmetical tasks. One may even hazard the guess that, with no other change in conditions than that of equalizing teaching time, the Cui. L subjects might have surpassed the Con. H subjects in conceptual maturity in all three tasks.

If one forgets for the time being the generally favorable position of the two H groups, one might conclude that over-all achievement in arithmetic was a real factor in determining the relative effectiveness of the two programs. But this conclusion is unwarranted if one holds in mind the probable influence of differences in other factors. Hence, the results of neither of the comparisons, (5) or (6), are helpful in answering question C.

Reading achievement. Neither do the results of the comparisons (5) and (6) for the factor reading achievement help to answer question C. In (5) the Cui. H subjects excelled in level of conceptual maturity in all three arithmetical tasks when compared with the Con. L subjects; but in (6) the Con. H subjects excelled not at all in types of thought processes used with the number combinations, excelled slightly in accounting mathematically for computations, and excelled by a comfortable margin only in ability to explain solutions of problems.

In (5) the Cui. H group, as usual, had the higher ratings in factors other than reading achievement, but had to overcome whatever handicap was imposed by having had 14% less teaching time than the Con. L subjects. In (6) the Con. H subjects were judged to be much brighter and much better readers than the Cui. L subjects and, besides, had had 26% more arithmetic instruction. In both (5) and (6) the possible influence of differences in reading achievement on the relative effectiveness of the two programs was almost certainly overshadowed by the influence of differences in other factors.

Additional data on the relative effectiveness of the programs

Certain data not examined heretofore throw new light on the relative effectiveness of the Cui. and of the Con. programs in promoting conceptual maturity.

In the study of each of the four factors, six pairs of groups have been compared with respect to level of conceptual maturity in dealing with each of three arithmetical tasks, - a total of eighteen comparisons for each factor and a total of 72 comparisons for the four factors together.

However, since in all comparisons of the types (1) and (2) both the groups compared were drawn from the same program sample (Cui. H vs Cui. L, and Con. H vs Con. L), the findings of the 24 (1) and (2) comparisons have nothing to do with the relative effectiveness of the two programs. They must therefore be eliminated in the present analysis, thus leaving a total of 48 comparisons that have at least potential value.

The twelve (3) to (6) comparisons for groups set up for reading achievement should also be omitted. Reading achievement is so closely related to brightness and over-all achievement in arithmetic that the findings of comparisons of reading groups can do little more than to duplicate those for the other two factors just named. Inclusion of the results for the reading groups conceivably could unjustifiably exaggerate the difference between the two programs in effectiveness, in favor of one program or the other.

So, in the end there remain 36 comparisons of Cui. and Con. groups which can legitimately be used for the present purpose.

As the reader may recall, in most pairs of groups the Con. subjects had been given more instruction than the Cui. subjects. In the following paragraphs such differences, as well as differences in other factors, will not be taken into consideration. This practice should introduce no injustice for or against either program, save in the case of time, for the advantages in other factors balance out fairly well for the Cui. and the Con. groups.

To return to the data: in 27 of the 36 comparisons a Cui. group or a Con. group -- whether H or L is of no importance here -- was slightly, somewhat, or substantially superior to the other group in the pair in level of conceptual maturity. In 24 of the 27 comparisons this distinction belongs to Cui. subjects and in three, to Con. subjects. Thus, the conclusion reached in Chapter 4 is confirmed: in the Scottish schools the Cui. program was much more effective than the Con. program in moving children steadily higher in the scale of meaningful arithmetical abstractions.

Conceptual maturity vs proficiency

The problem here may be stated in the form of the question, Is progress in conceptual maturity purchased at the expense of progress in proficiency? Children who are well advanced in the former may be lacking in proficiency, or they may be superior in proficiency.

Once again the reader is reminded that the term "proficiency" is used here in a very restricted sense. It does not refer to ability to find answers for all combinations, or to skill in computing answers for all types of examples taught, or to success in solving all one-step problems. Rather, it refers only to measures of accuracy and rate in connection with the small sample of items in each of the arithmetical tasks.

For each factor, six pairs of groups were compared, not only with respect to their relative conceptual maturity, but also with respect to their relative proficiency in each arithmetical task. Since for the present purpose there is no need to distinguish between Cui. and Con. groups, the results of comparisons (1) and (2) can be combined with those for (3) to (6). If the comparisons for the factor, Reading Achievement, are again omitted, for the reason given above, there are 54 comparisons by which the relationship between conceptual maturity and proficiency may be studied.

Ten of the 54 comparisons are lost because neither of the groups compared was superior to the other in conceptual maturity. In 24 of the remaining 44 comparisons, greater conceptual maturity is positively related to superiority in proficiency; in six it is negatively related; and in 14 it is related not at all to proficiency, for the groups compared are equal in this respect.

On the whole, there is little evidence to support the hypothesis that emphasis on conceptual development necessarily means a lowering of proficiency on the part of the children taught. At the same time, one cannot say, on the basis of the evidence above, that emphasis on conceptual development necessarily means improvement in proficiency, for this increase in proficiency did not occur in nineteen of the 44 comparisons. Among these nineteen comparisons there must have been instances in which effort to produce proficiency was deliberately deferred for a time, in order to work more intensively on engendering arithmetical meanings and understandings. If so, the expectation would be that in the end prior attention to the conceptual aspects of arithmetic would pay large dividends in increased proficiency; and there is reason to believe that, if proficiency were stressed later on, the expectation would be realized.

PART III

THE INVESTIGATION IN ENGLISH SCHOOLS

CHAPTER 6

GROSS COMPARISON OF THE EFFECTIVENESS OF THE CUI., THE CON., AND THE DIEN. PROGRAMS

Data on Schools and Subjects

In the investigation in English schools data were collected concerning the results of instruction according to three programs instead of two. To the Cui. and the Con. programs, studied in Scottish schools but taught also in English schools, a third, the Dienes (Dien.) program, was added. This last program was, at the time, taught only in English schools and in relatively few of them.

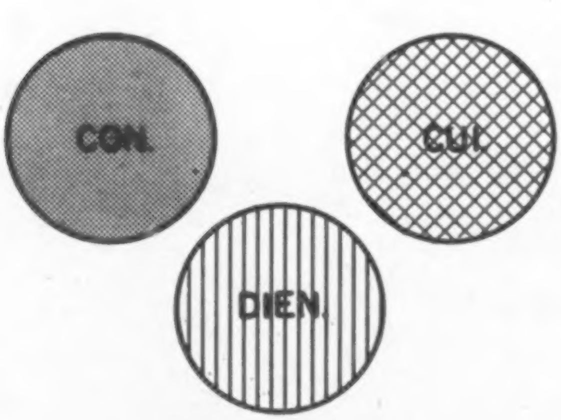

Fig. 7. Misrepresentation of the relationship between the Cui., Dien., and Con. programs.

The relationship among the three programs, in matters of objectives, content, and instructional materials, is not at all accurately pictured in Fig. 7, for it implies that each program is uniquely and completely distinct from the others. Actually, there is overlapping among them, as is depicted in Fig. 8, much more between the Dien. and the Con. programs than between the Dien. and the Cui. programs. For a fuller treatment of features common to, and different among the three programs, the reader is referred to Chapter 2.

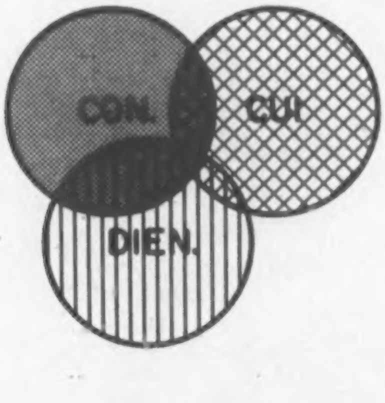

Fig. 8. Better representation of the relationship between the three programs.

The organization of Chapter 6, for the English study, will be the same as that of Chapter 4, for the Scottish study. The data procured in the Cui., the Dien., and the Con. schools in England will be treated first for the number combinations, then for the three examples for which answers were to be computed, and finally for the one-step problems which were to be solved; and for each of these types of arithmetical task the data relating to conceptual maturity will be presented prior to those for proficiency since the principal concern in the investigation as a whole is the relative effectiveness of the three programs in encouraging and directing children to move steadily to higher and higher levels of meaningful arithmetical abstractions.

Tables 1 to 11 for the Scottish subjects will be paralleled, topic by topic, by corresponding tables numbered 101 to 111 for the English subjects, in order to facilitate comparisons. Some slight modifications will be made in Tables 101 to 111 whenever they seem desirable.

Table 101 summarizes certain information concerning the cooperating English schools, of which there were 32 in all, furnishing a total of 928 subjects, for each of whom a complete interview schedule is available. The samples of schools and subjects are larger than was the case in the Scottish study; but there are the same wide ranges (a) in number of subjects supplied and (b) in time allotted to arithmetic by the schools in each group identified by the particular program in use. Fortunately, the weighted means for total instructional time are closely similar for the Cui., the Dien., and the Con. schools. That for the Con. schools for the three-year period is only 9% larger than that for the Cui. schools and only 7% larger than that for the Dien. schools.

The percentage distributions of the three groups with reference to sex, age, and teachers' estimates of brightness, general achievement in arithmetic, and reading achievement are reported in detail in Table 102. In none of the

Table 101. General Information Concerning Cooperating Schools

Type of Program	Number of Schools	Total Number of Subjects	Range in Number of Subjects	Time Allotment for Arithmetic, 3-Year Total	
				Range	Weighted Mean
Cui.	10	257	7 - 66	360 - 720	538
Dien.	10	358	3 - 45	360 - 720	545
Con.	12	313	7 - 75	410 - 800	585

groups is there a 50 - 50 balance between boys and girls, but also in none is the disproportion of one sex or the other large enough to be of any consequence in affecting the results of comparisons, group with group.

Because it is difficult to get a clear picture of the likeness of, and of the differences between the groups from the data for the other categories in Table 102, and because the groups will be compared as wholes, weighted means were computed according to the method explained in Chapter 4, and these are summarized in Table 103.

The Dien. subjects were, on the average, three months older than the Cui. and the Con. subjects; but this difference will be disregarded. The decision to do so could probably be justified purely on the ground that so small an age difference could scarcely have affected the relative success of the Dien. subjects in learning. On the other hand, this assumption could be unsound, and their being three months older may actually have made learning either easier or more difficult for them. Unfortunately, no means are ready to hand to determine

Table 102. Comparison of Characteristics of Cui., Dien., and Con. Subjects, in Per Cents

Characteristic		Per Cents, by Program		
		Cui.	Dien.	Con.
Sex	Male	47	51	52
	Female	53	49	48
Age	7 yr. 5 mo. or less	7	4	1
	7 yr. 6 mo. to 7 yr. 11 mo.	35	19	33
	8 yr. 0 mo. to 8 yr. 5 mo.	32	41	43
	8 yr. 6 mo. and over	26	37	23
Estimated Brightness	Rating 1. (Brightest)	34	24	35
	2.	29	28	27
	3.	28	32	27
	4.	9	15	12
Estimated General Achievement in Arithmetic	Rating 1. (Highest)	28	19	29
	2.	26	29	30
	3.	33	34	28
	4.	13	17	12
Estimated Reading Achievement	Rating 1. (Highest)	29	30	41
	2.	28	30	25
	3.	30	23	21
	4.	12	17	14

Table 103. Characteristics of Subjects, Weighted Means

Type of Program	Age	Rating in Estimates for		
		Brightness	Achievement in Arithmetic	Achievement in Reading
Cui.	7 yr. 10 mo.	2.12	2.31	2.23
Dien.	8 yr. 1 mo.	2.36	2.47	2.27
Con.	7 yr. 10 mo.	2.18	2.21	2.10

whether the age difference was a special advantage or disadvantage to the Dien. subjects, and if either, by how much.

As a whole, the ratings for other factors indicate that the Con. subjects, though slightly less bright than the Cui. subjects, seem to have been the ablest in school learning; and the Dien. subjects are rated as the least bright and the least capable in school learning. However, the differences in comparisons of ratings are not generally large. In only two of the nine comparisons does the difference amount to 0.24 or more of a rating, and in only three more does it amount to 0.18 of a rating.

The Number Combinations

Conceptual maturity: Level of thought processes

Tables 104 and 105 contain the data needed to compare the Cui., the Dien., and the Con. subjects with respect to their relative levels of conceptual maturity in ways of thinking answers for the number combinations. Table 105,

Table 104. The Number Combinations. Thought Processes with the Eight Combinations, by Rubrics; Averages in Per Cents

Rubrics	Cui.	Dien.	Con.
Indeterminate	0	1	1
A. Non-Attempts	9	12	3
B. Counting, All Forms	23	32	24
C. Guessing, Rote Recall, Recitation of Tables, etc.	26	21	22
D. Rational Solutions, All Forms	35	29	27
E. Uses of Meaningful Habituations	5	4	22

the more detailed of the two, catalogues, by categories and by rubrics and in terms of average per cents, the thought processes of the three groups for the pairs of combinations. Table 104, a modified form of Table 4 in the Scottish study, reports the same data as in Table 105, but for the eight combinations as a whole and classified by rubrics instead of categories.

All combinations. The reader will recall (a) that Rubrics A and B represent the lowest levels of thought processes considered from the viewpoint of conceptual maturity; (b) that Rubric C may or may not represent the next higher level, but that because of uncertainty at this point the per cents for the categories in this rubric are usually disregarded in comparisons; (c) that

Rubric D represents next to the highest level of conceptual maturity, but that in general the per cents for the categories in this rubric tend to decrease in size as the per cents for the most mature level increase; and (d) that Rubric E, Uses of Meaningful Habituations, represents the most mature level of thought process.

According to Table 104, the average per cents for Rubrics A and B combined are, in descending order, 44 for the Dien. group, 32 for the Cui. group, and 27 for the Con. group. Stated differently, the Con. subjects were slightly less inclined than were the Cui. subjects to react to the number combinations by processes at the lowest end of the conceptual scale, and they were significantly less inclined to do so than were the Dien. subjects.

The average per cents for Rubric E, at the other end of the maturity scale, are unexpectedly small,* the largest being 22 for the Con. subjects, and it is at least four times as large as that of each of the other two groups, which made use of Meaningful Habituations to about the same extent.

The average per cents for Rubric D show the trend mentioned in (c) in the second paragraph above: with the largest average per cent for Rubric E goes the smallest average per cent for Rubric D.

All in all, the Con. subjects exhibited, by comfortable margins, the highest degree of conceptual maturity of the three groups in thinking answers for the eight combinations, for they have the lowest average per cent for Rubrics A and B combined and by far the highest average per cent for Rubric E. Second to the Con. subjects are the Cui. subjects, by reason of their significantly smaller average per cent for Rubrics A and B combined, as compared with the Dien. subjects, and their slightly larger average per cent for Rubric D.

*** In the Scottish study, the average per cent for Rubric E is 36 for the Cui. group, while that for the Con. group is 15.**

Table 105. The Number Combinations. Thought Processes Operation by Operation; Averages in Per Cents

Rubrics and Categories	Operation											
	Addition			Subtraction			Multiplication			Division		
	Cui.	Dien.	Con.	Cui.	Dien.	Con.	Cui.	Dien.	Con.	Cui.	Dien.	Con.
Indeterminate	0	0	0	0	1	2	0	1	0	0	1	0
A. Non-Attempts	0	0	0	6	4	2	7	10	3	22	35	7
B. Counting, All Forms	46	61	45	32	50	39	12	9	5	2	6	5
C. Guessing	1	0	2	4	4	5	6	6	5	8	11	9
Rote Recall	3	2	2	2	2	2	13	8	1	6	8	3
Recitation of Tables	0	0	0	0	0	0	21	24	33	15	6	13
Mechanical Uses: Reverses and Inverses	2	0	2	7	2	2	3	2	1	14	10	13
Sub-Totals	6	2	4	13	8	9	43	40	40	43	35	38
D. Rational Uses: Reverses and Inverses	0	0	1	2	1	1	1	1	3	3	5	9
Solutions, Other Forms	37	30	38	40	33	42	31	30	9	24	17	7
Sub-Totals	37	30	39	42	34	43	32	31	12	27	22	16
E. Uses of Meaningful Habituations	7	4	10	4	2	5	4	9	40	5	1	34

Combinations, by operation. The reactions of the three groups to the addition and the subtraction pairs of combinations (Table 105) were quite similar. When subjects did not "know" answers, in the sense of being able to use Meaningful Habituations, and few did, they counted (32% to 61%) or they made use of some form of Rational Solution (30% to 43%).

With the pairs of multiplication and division pairs, the three groups also reacted quite similarly, but the pattern in this case is quite unlike that for the addition and the subtraction and multiplication pairs: Non-Attempts were commoner rather generally; Counting decreased; Rubric C responses (especially the Recitation of Tables) increased from an average of about 10% for the three groups for the addition and the subtraction pairs to an average of about 40% for the multiplication and the division pairs; the frequency of Rational Solutions declined somewhat; and the use made of Meaningful Habituations remained about the same, except for the Con. subjects.

The average per cent of the Dien. group for Rubrics A and B combined is significantly larger than that of the Con. group for all four pairs of combinations, and it is substantially larger than that of the Cui. group for all pairs except that in multiplication. The Con. group has a significantly smaller average per cent for these two Rubrics together than does the Cui. group only for the multiplication and division pairs.

Use of Meaningful Habituations was somewhat commoner among the Con. subjects than among the Dien. subjects with addition combinations, but was far commoner among the Con. subjects than among those in the other two groups with the multiplication and division combinations.

Rational Solutions were employed considerably less often by the Con. subjects than by the Cui. and the Dien. subjects in the case of the multiplication and the division pairs, but probably because of their far greater utilization of

Meaningful Habituations in these instances.

The study of thought processes with the pairs of combinations confirms the evaluation reached in the study for the eight combinations as a whole: (a) the clear superiority of the Con. group over each of the other two groups in conceptual maturity and (b) the slight superiority of the Cui. group over the Dien. group in this respect.

Proficiency: Accuracy and rate

Accuracy. According to the average per cents in the last column of Table 106, the Con. subjects attempted more answers for the eight combinations (96) than did the Cui. subjects (91) and the Dien. subjects (87), and they are credited with a relatively larger proportion of Correct Answers than is either of the two other groups, -- 79 compared with 74 (Cui.) and 69 (Dien.). While the differences between pairs of average per cents are small for each category, Non-Attempts and Correct Answers, the consistency of the results warrants the claim that the Con. group was at least somewhat more accurate than the Cui. group and was certainly more accurate than the Dien. group.

The comparative accuracy of the three groups for the four pairs of combinations classified by operation is as follows:

Addition: no difference.

Subtraction: Con. group slightly more accurate than the Cui. group and somewhat more accurate than the Dien. group; little difference between the Cui. and the Dien. groups.

Multiplication: Con. group significantly more accurate than both of the other groups; no difference between the Cui. and the Dien. groups.

Division: Con. group significantly more accurate than both of the other groups; Cui. group significantly more accurate than the Dien. group.

Even though its advantage is clearly marked in the case only of the multiplication and the division pairs, the Con. group should now be regarded as

Table 106. The Number Combinations. Proficiency in Giving Answers: Accuracy; Averages in Per Cents*

Programs and Measures	Operation				
	Addition	Subtraction	Multiplication	Division	Averages
Cuisinaire					
Correct Answers	91	71	69	67	74
Non-Attempts	0	7	6	22	9
Dienes					
Correct Answers	93	68	67	50	69
Non-Attempts	0	5	10	36	13
Conventional					
Correct Answers	92	75	79	72	79
Non-Attempts	0	3	3	9	4

* Per Cents for Incorrect Answers are omitted

having demonstrated considerable superiority in accuracy over each of the other two groups; and the Cui. group, somewhat greater accuracy than the Dien. group.

Rate. By itself, quickness of response is an untrustworthy index of proficiency, for it takes no more time to blurt out an incorrect answer, perhaps a blind guess, than it does to announce a correct answer. But promptness of response, when the reply is correct, probably attests confidence on the part of the learner, and confidence is an important dimension of real proficiency.

As shown by the entries in the last column of Table 107, the Con. subjects have a materially larger average per cent for Prompt Responses than does either

Table 107. The Number Combinations. Proficiency in Giving Answers: Rate; Averages in Per Cents*

Programs and Measures	Operation				
	Addition	Subtraction	Multiplication	Division	Averages
Cuisinaire					
Prompt Responses	39	22	29	38	32
Long Pauses	14	30	27	24	24
Dienes					
Prompt Responses	49	25	30	33	34
Long Pauses	8	31	30	38	27
Conventional					
Prompt Responses	50	29	47	46	43
Long Pauses	6	26	23	25	20

* Per Cents for Slight Pauses are omitted

of the other two groups, - this, for the eight combinations together. The Con. subjects also have the smallest average per cent for Long Pauses of the three groups, -- slightly smaller than the Cui. group and somewhat smaller than the Dien. group. The records both for Prompt Responses and for Long Pauses are about the same for the Cui. and the Dien. subjects.

For each pair of combinations, the Con. average per cent for Prompt Responses is the largest of the three, -- significantly so only once for the addition pair (in comparison with the Cui. group) and once for the multiplication pair (in comparison with the Dien. group), but significantly so for the division pair in both comparisons. Compared with the Dien. subjects, the Cui. subjects have a somewhat larger per cent for this category for one pair and the smaller average per cent for three pairs, the difference being significant only for one pair.

Summary: The number combinations

In ways of thinking answers for the number combinations the Con. subjects revealed a considerably higher degree of conceptual maturity than did the Cui. and the Dien. subjects, and the Cui. subjects revealed a slightly higher degree of conceptual maturity than did the Dien. subjects.

In proficiency also the Con. group greatly surpassed the other two groups, both an Accuracy of answers and in Rate of response, and the Cui. group surpassed the Dien. group by small margins.

Computation

Conceptual maturity: Ability to explain computations

Table 108 differs from Table 8 only in that the data for the category, Other, have been separated from those for the category, Imagery, and are reported, while the data for Imagery, used very rarely, are omitted.

The average per cents for the total of Satisfactory Explanations of the four computations combined are 36 for the Cui. subjects, 33 for the Dien. subjects, and 54 for the Con. subjects. The Cui. and the Dien. subjects were, then, about equal in ability to make intelligible their understanding of the mathematical rationale of the computations, but in only about a third of their attempts. The Con. subjects, on the other hand, were able to do so in more than a half of their attempts. In level of conceptual maturity as measured here the Con. subjects far outrank the Cui. and the Dien. subjects, between

Table 108. Computation. Quality of Explanations of Computations, Attempts Only: Per Cents and Average Per Cents*

Programs, and Quality of Explanation	Examples				
	27 + 46	92 - 45	28 x 3	4/84	Averages
Cuisinaire					
Satisfactory					
Words Only	4	2	7	21	8
Altered Examples	49	24	40	1	28
Other	0	1	0	0	0
Unsatisfactory					
Words Only	23	26	18	74	35
Altered Examples	22	40	29	2	23
Other	2	7	5	2	4
Dienes					
Satisfactory					
Words Only	24	14	14	30	20
Altered Examples	23	29	22	0	18
Other	3	3	4	1	3
Unsatisfactory					
Words Only	26	34	25	55	35
Altered Examples	12	8	16	0	9
Other	10	12	19	12	13
Conventional					
Satisfactory					
Words Only	12	6	8	54	20
Altered Examples	48	47	43	0	34
Other	0	0	0	0	0
Unsatisfactory					
Words Only	11	16	17	43	22
Altered Examples	27	30	30	2	22
Other	1	1	2	1	1

* Per cents for Imagery are omitted

whom there is no difference in this regard.

Credit was given for Satisfactory Explanations regardless of the method employed, whether expressed in terms of place value or otherwise. It had been expected that the Dien. subjects would surpass the Cui. and the Con. subjects in relative success in accounting mathematically for computations. They had used the M. A. B. (see Chapter 2) extensively as concrete aids, and these materials would appear to be especially helpful in developing understanding of the structure of numbers and the process of renaming numbers in computation.

Yet, the expected did not occur. While the Dien. subjects used place value in many of their explanations, they also used methods classifiable only as Other more extensively than did the Cui. and the Con. subjects, and not very effectively at that. The Dien. average per cent for the total of explanations, acceptable and unacceptable, in the category Other is 16, to be compared with 4 for the Cui. subjects and 1 for the Con. subjects. And the Dien. average per cent of Satisfactory Explanations in the category Other is 3. All this means that only a fifth of the Dien. Other attempts were regarded by interviewers as Satisfactory.*

---

* It may occur to the reader that the interviewers of Dien. subjects were unacquainted with the relatively new M. A. B. and were therefore unqualified to interpret correctly the explanations offered by Dien. children. Such was not the case, however. A large part of the sample of Dien. subjects were interviewed by headmasters in schools in which instruction was in accordance with the Dien. program; and the other interviewers were familiar with the M. A. B. to various extents.

As a matter of fact, the four Dien. headmasters interviewed children in schools where the Dien. program was not taught. And some interviewers of Cui. subjects, though they were familiar "to various extents" with the Cui. program, were more at home with the Con. program, while others who interviewed Con. subjects knew the Con. program less intimately than they did, say, the Cui. program.

The practice of having interviewers work with children taught according to at least two of the programs whenever possible (and it was practically always possible) was deliberately adopted as one means of counteracting the effects of possible prejudice for a particular program, or at least the effects of greater knowledge of one program than of the others.

To turn now to the data for the separate computations: the average per cents of the three groups for totals of Satisfactory Explanations are as follows, in the order Cui., Dien., Con.:

27 + 46	53; 50; 60	3 x 28	47; 40; 51
92 - 45	27; 46; 53	84 ÷ 4	22; 31; 54

Without exception the Con. group offered more Satisfactory Explanations than did either of the other two groups for each computation. Compared with the Cui. and the Dien. average per cents for this category, those of the Con. group are enough larger to make differences significant or nearly so in five of the eight chances. The record of the Cui. subjects about equals that of the Dien. subjects for one example, is better for one example, and is inferior for two examples.

These comparisons, example by example, do not alter the ranking of the three groups in conceptual maturity as determined by the comparisons above for the four examples as a whole, except possibly to indicate that the Dien. group may have been very slightly superior to the Cui. group.

Proficiency: Accuracy and quality of performance

Accuracy. The Con. group was considerably more skillful than each of the other groups in the four computations. The data to support this conclusion are to be found in the last column of Table 109.

To consider first the data for the category Correct Answers: the average per cent of the Con. subjects is considerably larger than are those of the Cui. and the Dien. groups, and the average per cent of the Cui. group is considerably larger than is that of the Dien. group.

The relative standing of the three groups with respect to frequency of Non-Attempts is the same as in the comparisons with respect to Correct Answers, the difference in average per cents in each comparison being significant or nearly so (9 percentage points in the Cui. vs Con. comparison).

Table 109. Computation. Proficiency: Accuracy and Quality of Performance: Per Cents and Average Per Cents

Programs, and Types of Measure	Examples				Averages
	27 + 46	92 - 45	28 x 3	4/$\overline{84}$	
Cuisinaire					
Accuracy*					
Correct Answers	72	54	51	69	61
Non-Attempts	3	11	11	22	12
Quality of Performance**					
Smooth	81	72	68	90	78
Erratic	19	28	32	11	22
Dienes					
Accuracy*					
Correct Answers	65	38	35	46	46
Non-Attempts	10	24	30	42	26
Quality of Performance**					
Smooth	58	47	40	65	52
Erratic	42	53	60	35	47
Conventional					
Accuracy*					
Correct Answers	85	71	65	81	75
Non-Attempts	1	2	3	8	3
Quality of Performance**					
Smooth	85	73	72	89	80
Erratic	15	27	28	11	20

* Per cents for Incorrect Answers are omitted
** Attempts only

In credit for Smooth Performances, the Cui. and the Con. groups were a match for each other, but both groups surpassed the Dien. group by wide margins.

The Con. subjects made much the best showing of the three groups in the category Correct Answers for each of the four examples, and the Cui. subjects made a much better showing than did the Dien. subjects in all four examples.

The Dien. subjects consistently have the highest average per cent for Non-Attempts; and in all eight comparisons their average per cents are smaller by 9 percentage points or more. Three of the four Cui. vs Con. comparisons favor the Con. group. In two, the difference between average per cents is 8 or 9 percentage points; in the third, 14 percentage points.

In every example the Dien. group has by far the smallest per cent for Smooth Performances. The Cui. group has a larger per cent than has the Con. group for one example and a smaller per cent for another. So, there is little to choose between the Cui. and the Con. groups in this category.

Summary: Computation

In quality of explanations of computations the Con. group demonstrated a degree of conceptual maturity far in advance of that of the Cui. and the Dien. groups, and the last-named two groups were apparently about equally matched.

In computational proficiency the Con. group again demonstrated unmistakable superiority over each of the other two groups, and the Cui. group, over the Dien. group.

Problem Solving

Conceptual maturity: Ability to explain solutions

The three groups were remarkably alike in ability to account for solutions of the twelve problems. For each of the four categories in Table 110 comparisons

Table 110. Problem Solving. Reasons Given for Solutions by Operation, Attempts Only; Averages in Per Cents

Programs, and Quality of Reasons	Problems, by Operation				
	Addition	Subtraction	Multiplication	Division	Averages
Cuisinaire					
Good	65	46	56	36	51
Fair	24	28	24	30	26
Unsatisfactory	8	21	16	30	19
Indeterminate	3	5	3	4	4
Dienes					
Good	76	51	46	37	52
Fair	19	29	30	32	27
Unsatisfactory	3	17	21	27	17
Indeterminate	2	3	3	4	3
Conventional					
Good	65	48	57	40	52
Fair	17	28	22	37	26
Unsatisfactory	11	15	14	16	14
Indeterminate	7	8	7	7	7

can be made for three pairs of groups: Cui. vs Dien., Cui. vs Con., and Dien. vs Con. The largest difference in the three pairs of average per cents is 1 point for the categories Good and Fair Explanations, is 5 points for Unsatisfactory Explanations, and is 3 points for Indeterminate Explanations.

The Dien. group was, by substantial margins, the most successful of the three in explaining the addition problems. All three groups have about the same average per cent for Good Explanations for the division problems, and the Cui. and the Con. groups, for the other sets.

The record made by the Dien. subjects was unexpectedly good. Their level of conceptual maturity was much lower than that of the Con. subjects both in work with the number combinations and in accounting mathematically for the four computations. Yet, in ability to explain solutions of the interview problems they equalled the Con. subjects in conceptual maturity.

If one were to hazard an explanation, it probably could not be found in the special activities of the Dien. subjects in manipulating the M. A. B., for these activities relatively seldom were related to quantitative problems imbedded in social situations. Rather, it might be found in their many experiences in solving practical, realistic problems having their origin in every-day life situations in and out of school.

In his account of the arithmetic program in the Leicestershire schools (Chapter 2), Mr. Sealey makes it clear that the problem solving experiences of the Dien. subjects were specifically designed to supplement work with the M. A. B. For this reason the unexpected success of the Dien. subjects in solving problems in the interviews can scarcely be attributed directly to the one feature which more than any other in the first three school years differentiates the Dien. program from the Cui. and the Con. programs.

This is not to say, however, that the many experiences of the Dien. subjects with the M. A. B. had no beneficial effects whatsoever on their ability to explain solutions of problems. On the contrary, they may have contributed indirectly by affording opportunities to gain insight into the nature of the basic operations and to increase understanding of number relationships, and the like;

Table 111. Problem Solving. Proficiency: Accuracy and Performance Level, by Operation; Averages in Per Cents

Programs, and Types of Measure	Problems, by Operation				
	Addition	Subtraction	Multiplication	Division	Averages
Cuisinaire					
Accuracy*					
Correct Answers	89	79	76	71	79
Non-Attempts	2	4	5	9	5
Performance Level**					
High	55	30	53	29	42
Low	23	37	24	24	27
Dienes					
Accuracy*					
Correct Answers	93	80	86	70	82
Non-Attempts	1	4	4	12	5
Performance Level**					
High	61	39	41	17	39
Low	16	30	22	30	24
Conventional					
Accuracy*					
Correct Answers	94	83	82	79	84
Non-Attempts	0	2	1	5	2
Performance Level**					
High	79	53	64	47	61
Low	10	23	17	18	17

* Per cents for Incorrect Answers are omitted
** Per cents for Intermediate Performance Level are omitted

and such increments of mathematical knowledge the Dien. subjects could have transferred to problem solving. On the whole, however, the major contribution was more probably made by the unusually vital character of the instruction offered in the Dien. schools with respect to problem solving.

Proficiency: Accuracy and performance level

Inspection of the average per cents of the three groups for Accuracy in Table 111 shows: for the category Correct Answers, no difference in the three comparisons larger than 5 points; for the category Non-Attempts, no difference larger than 3 points. With regard to Performance Level, two of the three comparisons for the category High-Level Performances show significant differences, both favoring the Con. group; and for the category Low-Level Performances, only one of the three comparisons shows a significant difference, this too in favor of the Con. group. Since Accuracy is a more critical measure of proficiency than is Performance Level, the only possible conclusion is that the three groups equalled each other in skill in solving the twelve problems.

The results of the three groups with regard to proficiency in the four sets of problems substantiate the assessment just made for the twelve problems as a whole. For no set of problems is there a significant difference between the average per cents in any of the three pairs for either of the categories Correct Answers and Non-Attempts. The Con. group is consistently credited with more High-Level Performances than is either of the other two groups for all four sets, significantly so in all of the twelve possible comparisons. In the same category the Cui. group surpassed the Dien. group by considerable margins for two sets, and the Dien. surpassed the Cui. group significantly in none. The data for the category Low-Level Performances add nothing to the general picture.

Summary: Problem solving

The three groups were a match for each other in ability both to explain solutions of problems and to find correct answers for problems.

Now, all twelve problems required knowledge of number combinations, in the sense of being able somehow to find their correct answers. Yet, according to Table 106 the three groups were unequally proficient in doing so for the eight interview combinations. The Dien. group was much less proficient with these combinations than was the Con. group and somewhat less proficient than was the Cui. group. Why, then, was the Dien. group as successful as were the Con. and the Cui. groups in offering correct answers for the one-step problems?

Two explanations (and there may be more) come to mind. (a) The number combinations in the problems may have been "easier" than those in the interviews so that the children in all three groups were readier with answers for the problem combinations. There is some ground for the belief that this may have been true, as may be inferred from a comparison of the two sets of number combinations.

Interview combinations	8 + 5	15 - 7	6 x 4	21 ÷ 7
	6 + 9	17 - 8	3 x 9	32 ÷ 8
Problem combinations	5 + 6	16 - 7	5 x 5	14 ÷ 4
	9 + 7	14 - 9	3 x 4	20 ÷ 4
	5 + 5	12 - 8	5 x 6	18 ÷ 6

(b) On the other hand, getting correct answers for the one-step problems involves more than mastery of the number combinations. Among other things, it practically always requires selection of the suitable operation, which in turn is dependent upon the ability to recognize in the varied language of problems the need for this or that operation (as it is not in abstract examples, in which there are a few relatively simple cues, easily learned). It has been shown that

the Dien. subjects, lowest of the groups in mastery of the interview combinations, were equal to the other two groups in ability to explain solutions of problems. Ability to account rationally for solutions implies, in large measure, understanding of the function of each basic operation and ability to identify in one-step problems the appropriateness or inappropriateness of its use. These perquisites to proficiency in problem solving the Dien. subjects had apparently acquired, to an extent that balanced whatever their relative shortage in knowledge of the abstract number combinations.

There is no way to ascertain whether explanation (a) is more or less valid than is explanation (b). The fact is that the best explanation of the unlooked-for proficiency of the Dien. subjects may be a combination of (a) and (b), -- or indeed it may be something else entirely.

Chapter Summary

Conceptual maturity

This chapter reports the results of an investigation to determine the relative effectiveness of the Cui., the Con., and the Dien. programs, as taught in English schools, in promoting conceptual maturity in three types of arithmetical task. Comparisons were made among the three total samples of subjects, numbering in excess of 250 in each program group, drawn from at least ten different schools in each case. The findings are:

(a) The Con. program was far more effective than was each of the other two programs, and the Cui. program but slightly more effective than was the Dien. program in moving children to steadily more mature (meaningfully abstract) ways of thinking answers for the eight interview combinations, and supposedly in thinking answers for all simple number combinations.

(b) The Con. program was far more effective than was each of the other two programs in developing ability to explain the mathematical rationale of the computations for the four interview examples, and supposedly of the computations of all examples that had been taught. The Cui. and the Dien. programs, much less successful in this respect than the Con. program, were equally effective.

(c) The three programs were equally effective in teaching children how to explain their solutions of the twelve one-step interview problems, and supposedly of all such problems.

The superior effectiveness of the Con. program in (a) and (b) can hardly be explained as owing to the fact that the Con. subjects had had more instruction in arithmetic than the Cui. and the Dien. subjects had had. Distributed over three years, the Con. subjects were taught not more than ten to twenty minutes a week more than were the other two groups of subjects; and this difference in time would seem to be too small to have had much, if any, influence on the results of the comparisons.

The Con. group was estimated, on the average, to be very slightly less bright than the Cui. group, by 0.06 of a rating, and slightly brighter than the Dien. group, by 0.18 of a rating. According to estimates, the Con. subjects, on the average, were thought to have demonstrated somewhat greater success in school learning: in arithmetic achievement, by 0.10 of a rating over the Cui. subjects and by 0.26 of a rating over the Dien. subjects; in reading achievement, by 0.13 of a rating over the Cui. subjects and by 0.17 of a rating over the Dien. subjects. If the differences had amounted to, say, 0.40 or 0.50 of a rating (the largest is 0.26), the advantages of the Con. subjects in ability in academic learning would have had to be taken seriously into account in explaining their better showing in conceptual maturity in two of the three kinds of arithmetical task; but they would also raise the question, Why not also in the third kind as well?

It is evident that there is a distinct conflict in the relative effectiveness of the Cui. and the Con. programs in the English and in the Scottish schools, in which the Cui. program had much the better of the argument. More will be said about this matter in Chapter 8. Meanwhile, it proves to have been wise to have cautioned the reader in the closing section of Chapter 4 (which was actually written before the English data had been analyzed) not to infer too much from the superior conceptual maturity of the Cui. subjects in the Scottish schools. He was warned that the Cui. program per se had not been proved more effective than the Con. program and was advised to wait for the results in English schools before arriving at a final judgment in this matter.

One last bit of evidence, based on Table 108, to underscore the greater degree of conceptual maturity of the English Con. subjects in comparison with the English Cui. subjects: On the average, the Cui. subjects used crutches in computing answers for 51% of their examples, and they were successful in 9 out of 20 of their attempts to explain the computations in the examples so altered. The corresponding figures for the Con. group are 56% and 6 out of 10. In other words, crutches appear to have contributed more to understanding on the part of the Con. subjects and more to mechanical computation on the part of the Cui. subjects.

The Cui. subjects tried to explain their computations with Words Only, on the average, for 43% of their examples and were successful in 2 out of 20 attempts. The figures for the Con. subjects are 42% and 5 out of 10. When the Con. subjects relied on purely verbal explanations, a method higher in terms of conceptual maturity than the use of Altered Examples, they knew better than the Cui. subjects what they were talking about.

Proficiency

Thus far nothing has been said, by way of summary, about the relative effectiveness of the three programs in developing proficiency; that is to say, proficiency limited to work in connection with the particular interview items in the three arithmetical tasks.

(a) Number combinations. The Con. subjects were considerably more accurate and much quicker in response than were either the Cui. or the Dien. subjects, and the Cui. subjects were, by a small margin, more proficient than were the Dien. subjects.

(b) Computation. The Con. subjects made by far the best of the three records both for Accuracy and for Quality of Performance, and the Cui. subjects surpassed the Dien. subjects in proficiency by comfortable margins.

(c) Problem solving. The three groups were equally skillful in solving the twelve problems as far as Accuracy is concerned, with minor and somewhat conflicting differences as far as Performance Level is concerned.

In a word, in the two arithmetical tasks in which there are real differences among the groups in conceptual maturity, the group ranking highest in this dimension of learning (the Con. group) rated highest also in proficiency. In the third task, problem solving, no such comparison is possible, for the groups were equally matched in conceptual maturity.

CHAPTER 7

FACTORS POSSIBLY INFLUENCING THE EFFECTIVENESS AND THE RELATIVE EFFECTIVENESS OF THE PROGRAMS

This chapter, for the English study, serves the same purpose as did Chapter 5 for the Scottish study. In Chapter 5 pairs of groups differing in certain factors were compared to determine the influence of these factors on learning and, therefore, on the evaluation of the programs for effectiveness in promoting conceptual maturity. Five factors were identified: amount of instruction, brightness, arithmetic achievement, reading achievement, and age. The nature of the data collected for the age of subjects was such as to preclude the possibility of a fruitful investigation of this factor; and the findings respecting reading achievement were minimized since reading achievement is so closely associated with the factors, brightness and arithmetic achievement.

In this chapter the effects of differences in age will again have to be passed over, and for the same reason; and the effects of differences in reading achievement will not be discussed. Instead, only the basic tables for the latter factor will be supplied, and they will be transferred without comment to the Appendix. Full treatment will be accorded to the other three factors, and in the same order as in Chapter 5.

The organization of Chapter 7 will otherwise be about like that of Chapter 5. One dissimilarity is introduced by the fact that in the English study three programs, instead of two, must be compared for effectiveness, and there will be other minor dissimilarities. The tables will be numbered 113 to 132 to parallel Tables 13 to 32 in Chapter 5, and other tables will be added as needed, some designated as 118A, 118B, and so on, and some not numbered at all.

Like Chapter 5, Chapter 7 is long, and it is replete with tables and quantitative analyses. On this account the reader who is more interested in the findings than in the supporting evidence may wish to turn to the summaries, of which there is one for the chapter as a whole and one at the end of each of Sections 1 to 3.

1. Amount of Instruction as a Factor

The attempt to secure H and L groups with substantial differences in the total amount of arithmetic instruction they had received was much more successful in the English study than in the Scottish study. As shown in Table 112 for the English study, the H group, compared with the L group, had had 80% more instruction in the selected Cui. and Dien. schools and 52% more in the selected Con. schools. The numbers of subjects are adequate for comparisons in all of the six groups except in the Cui. L group. The percentage distribution by sex is odd in the two Dien. groups only. Differences appear in other factors, but these differences will be cited, and their effects on the results will be considered only in the final pages of this section, under the heading, Summary, Section 1.

A. The influence of wide differences in time upon the effectiveness of each separate program.

The problem here can be clarified by illustration. In some of the English Cui. schools arithmetic was allotted a relatively large amount of time over the three-year period; in others, a relatively small amount of time. One would assume that in the first-named schools (the H schools) the Cui. program would be more effective than in the second-named schools (the L schools); that is to say, that the Cui. program would carry children to greater heights of conceptual maturity in the H schools than in the L schools.

The relative success in the learning of the H and L groups is to be compared for three pairs of such groups: (1) Cui. H vs Cui. L, (2) Dien. H vs Dien. L, and (3) Con. H vs Con. L. In each of the three comparisons the relative conceptual maturity of the H and L groups (and, of secondary importance, their

Table 112. Time as a Factor. Composition of the H and the L Groups in Each Program

Bases of Comparison	Cuisinaire		Dienes		Conventional	
	H	L	H	L	H	L
Number of Subjects	93	36	120	88	106	97
Per Cents by Sex						
Male	42	58	63	37	55	48
Female	58	42	37	63	45	52
Weighted Means for						
Total Instructional Time	720	400	764	424	736	483
Age	7 yr. 10 mo.	7 yr. 6 mo.	8 yr. 1 mo.	8 yr. 2 mo.	7 yr. 11 mo.	8 yr. 0 mo.
Estimated Ratings:						
Brightness	2.32	2.08	2.35	2.54	1.98	2.30
General Achievement in Arithmetic	2.35	2.11	2.38	2.25	1.94	2.29
Achievement in Reading	2.44	2.19	2.65	2.56	2.02	2.40

relative proficiency) will be examined with reference to each of the three types of arithmetical task. This examination will be made purely in terms of the quantitative data for measures of learning as these data were assembled in tables, the more detailed of which are not reproduced here.

(1) Cui. H vs Cui. L

Conceptual maturity. (a) Despite having had only about 55% as much teaching time, the L subjects thought answers for the eight number combinations just about as did the H subjects. (Table 113.) For Rubrics A and B combined the average per cents are 35 for the L subjects and 41 for the H subjects; and the L average per cent for Rubric D is 4 points larger than that of the H subjects. Neither group used Meaningful Habituations at all commonly, but the H group did so slightly more often than did the L group.

For the four pairs of combinations, the H and the L groups compare as below. The first row in the tabulation is read: For the addition pair, the L subjects have a significant difference in their favor over the H subjects in Rubrics A and B combined (the smaller average per cent) and in Rubric D (the larger average per cent), while the H subjects have somewhat of an advantage in Rubric E, as is indicated by the parentheses in (H). A dash denotes no difference or one so small that it can be ignored.

Pairs of Combinations	Rubrics		
	A + B	D	E
Addition	L	L	(H)
Subtraction	--	(H)	--
Multiplication	L	--	--
Division	--	(L)	(L)

For none of the pairs does either group consistently and substantially surpass the other in level of abstractness of the thought processes used. On the contrary, the bulk of the evidence is that the two groups exhibited about the same degree of conceptual maturity.

Table 113. The Number Combinations. Thought Processes with the Eight Combinations; Averages in Per Cents

Rubrics and Categories*	Cuisinaire		Dienes		Conventional	
	H	L	H	L	H	L
A. Non-Attempts	11	11	10	2	2	2
B. Counting, All Forms	30	24	28	39	26	19
C. Guessing	6	6	6	6	5	8
Rote Recall	3	4	8	7	1	3
Recitation of Tables	9	15	6	14	9	13
Mechanical Uses: Reverses and Inverses	5	4	5	1	1	7
Sub-Totals	23	29	25	28	16	31
D. Rational Uses:						
Reverses	1	0	0	1	2	0
Inverses	2	2	3	1	1	2
Inverses, Plus	9	10	5	7	4	3
Solutions, Other Forms	17	21	23	14	21	23
Sub-Totals	29	33	31	23	28	28
E. Uses of Meaningful Habituations	7	1	3	5	27	19

* The categories Other, Imagery, and Indeterminate are omitted. The largest average per cent for Indeterminate is 0.8

Table 114. Computation. Quality of Explanations of Computations, Attempts Only; Averages in Per Cents

Classification of Explanations*	Cuisinaire		Dienes		Conventional	
	H	L	H	L	H	L
Satisfactory						
Words Only	9	4	13	26	28	21
Altered Examples	37	17	28	17	38	47
Other	0	0	3	2	0	0
Unsatisfactory						
Words Only	25	61	30	42	17	14
Altered Examples	18	17	14	5	17	16
Other	9	1	11	7	0	1

* The category Imagery is omitted; the average per cent in this category amounts to no more than 1.0 for any of the six groups.

(b) The situation with regard to ability to account mathematically for computations is quite different. The average per cents for Satisfactory Explanations of the four computations are 46 for the H group and 21 for the L group; and the difference between the H and L per cents for this category amount to 11 points or more for each of the four computations, all in favor of the H subjects, who in this particular arithmetical task were clearly more advanced conceptually than were the L subjects.

(c) The relative standing of the two groups in conceptual maturity is almost completely reversed in comparisons for ability to tell why problems were solved as they were. (Table 115.) Here the average per cent of the L subjects for Good

Table 115. Problem Solving. Reasons Given
for Solutions, Attempts Only; Averages in Per Cents

Quality of Explanations*	Cuisinaire		Dienes		Conventional	
	H	L	H	L	H	L
Good	39	48	50	67	75	49
Unsatisfactory	16	21	14	17	9	13
Indeterminate	5	11	4	1	5	4

* The category Imagery is omitted

Explanations of the twelve solutions exceeds the corresponding average per cent of the H subjects by 9 points. Moreover, the L average per cents for Good Explanations exceed those of the H subjects by 10 points or more for three sets of problems, while the H subjects were more successful, by 11 points, for one set only (division).

Proficiency. (a) The H group excelled in Accuracy of work with the eight combinations as a whole, especially in the category Correct Answers (Table 116), and also with the pairs of combinations. For two pairs its average per cents are the larger by 7 and 8 points; for the other two pairs, by 12 and 15 points.

The H subjects were inferior in Rate of response for the eight combinations, as shown in Table 116, and for the pairs of combinations. The L average per cent for Prompt Responses is much the larger for each pair. However, since greater emphasis is to be placed on Accuracy than on Rate, the H group is to be regarded as the more proficient in this arithmetical task, though its superiority is slightly lessened by its poorer record in Rate.

Table 116. The Number Combinations. Proficiency in Giving Answers: Accuracy and Rate; Averages in Per Cents

Nature of the Measures	Cuisinaire		Dienes		Conventional	
	H	L	H	L	H	L
Accuracy						
Correct Answers	73	62	68	78	80	83
Non-Attempts	11	16	11	3	3	3
Rate, Attempts Only						
Prompt Responses	31	52	44	22	47	44
Long Pauses	31	17	16	35	26	13

(b) The H subjects obtained significantly more Correct Answers than did the L subjects in working the four examples, but their claim to greater over-all Accuracy is diminished by their tendency to refuse to attempt more computations. (Table 117.)

While the H subjects have the larger per cent for Correct Answers in each separate computation, their advantage is significant only in two; and, at the same time, they were guilty of more Non-Attempts with each example, by significant amounts with two.

In Quality of Performance the H and L groups were about equal for the four examples combined. The H group worked considerably more smoothly than the L group in one computation, and the L group did so in two computations. On the whole, the H group is to be credited with but slightly more computational skill than is the L group.

(c) In solving the twelve problems the H subjects demonstrated somewhat greater Accuracy than did the L subjects, though there is nothing to choose between

Table 117. Computation. Proficiency: Accuracy and Quality of Performance; Averages in Per Cents

Types of Measure	Cuisinaire		Dienes		Conventional	
	H	L	H	L	H	L
Accuracy						
Correct Answers	55	44	52	63	82	79
Non-Attempts	19	9	15	13	1	3
Quality of Performance, Attempts Only						
Smooth	80	75	53	62	89	78
Erratic	20	25	47	38	11	22

Table 118. Problem Solving. Proficiency: Accuracy and Performance Level; Averages in Per Cents

Types of Measure	Cuisinaire		Dienes		Conventional	
	H	L	H	L	H	L
Accuracy						
Correct Answers	74	67	81	87	87	86
Non-Attempts	4	10	4	2	4	1
Level of Performance, Attempts Only						
High	45	43	38	41	80	61
Low	33	35	21	29	6	14

them in the matter of Level of Performance. (Table 118.)

The average per cents of the H group for Correct Answers are consistently the larger for the four sets of problems, but the difference between the two average per cents is significant for only one set (subtraction). The H group has the smaller average per cent for Non-Attempts for three sets, significantly so for only one set (again, subtraction). In only one comparison, for the division set, is there a real difference between the two groups in the category High-Level Performances, and this difference, in favor of the H subjects, is so substantial that it accounts for their having an average for this category for the twelve problems even as large as that of the L subjects.

The weight of the evidence is, then, that the H group was only slightly more proficient than was the L group in problem solving.

(2) Dien. H vs Dien. L

Conceptual maturity. (a) In ways of thinking answers for the eight combinations as a whole, the H and the L groups revealed similar degrees of conceptual maturity. (Table 113.) While the H subjects were somewhat more inclined to attempt answers than were the L subjects, they were less inclined to count. Hence, the two average per cents for Rubrics A and B together, rated lowest in the conceptual scale, are within 4 points of each other. Neither of the groups made more than slight use of Meaningful Habituations; but the H group did secure somewhat more answers by employing Rational Solutions than did the L group.

The amounts of difference, if any, between the average per cents of the two groups for the critical rubrics are summarized below, - this, for the pairs of combinations.

Pairs of Combinations	Rubrics		
	A + B	D	E
Addition	(H)	H	-
Subtraction	(H)	H	-
Multiplication	-	H	-
Division	(L)	(H)	-

While three of the seven advantages of the H subjects are not large enough to be significant, the consistency of the differences in their favor, particularly in Rubric D, argues for their having been slightly to somewhat more advanced in conceptual maturity than were the L subjects.

(b) One group was as successful as the other in accounting mathematically for the four computations as a whole, for their average per cents for Satisfactory Explanations of all kinds are 44 for the H group and 45 for the L group. (Table 114.) In these explanations the L subjects used Words Only significantly oftener than did the H subjects, and the latter subjects used Altered Examples significantly more commonly.

For only one of the four separate computations is there a real difference between the per cents of the two groups for Satisfactory Explanations, this one to the advantage of the L subjects in the subtraction example; but the same subjects also have an advantage of 7 points in the comparisons for the division example.

The category Other, omitted in Table 14 for the Scottish study, had to be included in Table 114 because it was needed to report the full record, primarily of the Dien. subjects. Attempts made by the Dien. subjects to explain computations, classified as Other, occurred in connection with all four examples, most commonly with the multiplication example and least commonly with the division example; and the totals of Other Explanations represent 14% of the H attempts and 9% of the L attempts. Presumably in their explanations the Dien. children were drawing upon what they had learned with the M. A. B., but not too successfully, for, regardless of the particular operation involved in examples, only 1 in 5 of the attempts to explain in each of the two groups, was taken to be adequate. Either of the Dien. subjects in question had not really come to a sound understanding of the mathematics of computation, or they were unable to verbalize their understanding by the method they chose for their explanations.

(c) According to Table 115, the L group exhibited far more ability than did the H group to explain their solutions of the twelve problems. And the L average per cent for Good Explanations exceeds that of the H subjects by 8 points for the division problems, and by 17 points or more for each of the other three sets of problems.

Proficiency. (a) The L group gave relatively more Correct Answers for the eight combinations than did the H subjects and tended somewhat less commonly to refuse to attempt answers. (Table 116.) The differences in average per cents for these two categories amount to 10 points and 8 points, respectively. The L group has the larger average per cent for Correct Answers for each of the four pairs of combinations, but significantly so for only two pairs.

On the other hand, the H group has much the larger average per cent for Prompt Responses and much the smaller average per cent for Long Pauses for the eight combinations as a whole and for each pair of combinations. Once more to assign priority to Accuracy, the L subjects surpassed the H subjects in proficiency in working with the number combinations.

(b) While both groups were disinclined, to about the same extent, to try to compute answers for the four examples, the L subjects significantly surpassed the H subjects in relative frequency of Correct Answers (Table 117), as they did also for three of the four separate computations. As for Quality of Performance, the L subjects worked smoothly and steadily somewhat oftener than did the H subjects in making the four computations, and they have the larger per cent for Smooth Performances for each example, with significant differences in two.

The comparisons of the two Dien. groups clearly show the L group to have been superior to the H group in computational proficiency.

(c) In skill in solving the twelve problems one group was a match for the other, both in Accuracy and in Level of Performance. None of the differences between the average per cents of the H and the L subjects in Table 118 is larger than 8 points for any of the four categories.

The H average per cent for Correct Answers is larger by 12 points than that of the L subjects for one problem set and smaller by 9 points for one set. The H subjects have a significantly larger average per cent for High-Level Performances for one set; the L subjects, for two sets.

(3) Con. H vs Con. L

Conceptual maturity. (a) To get answers for the eight combinations, the L group employed thought processes at the lower end of the conceptual scale slightly less often than did the H group (a percentage difference of 7 points for Rubrics A and B combined); but, on the other hand, the H subjects made somewhat greater use of Meaningful Habituations (a percentage difference of 8 points for Rubric E). The two average per cents for Rational Solutions are identical, (Table 113.) Despite a 15-point difference between the two average per cents for Rubric C, the records for this rubric had best be disregarded here as usual.

For the pairs of combinations the average per cents of the two groups differ in the selected rubrics as follows.

Pairs of Combinations	Rubrics		
	A + B	D	E
Addition	(L)	-	-
Subtraction	L	L	-
Multiplication	-	(H)	(H)
Division	-	L	H

The L group seems to have reacted to the subtraction pair with somewhat the greater degree of conceptual maturity; and the H group, to the multiplication pair. Again, the differences cancel each other. The safest judgment from all the data presented is that the two Con. groups had arrived at about the same level of abstractness of thought in obtaining answers for the number combinations.

(b) The two groups were equally able to explain the mathematical rationale of the four computations taken together, the only differences being in the preference for explanations with Words Only or with the aid of Altered Examples. (Table 114.) For two of the computations the H group has the larger per cent for Satisfactory Explanations, in neither by a significant amount; and for two the L group has the larger per cent for this category, one of them significantly so.

(c) The H subjects greatly surpassed the L subjects in ability to explain solutions for the twelve problems (Table 115), as they did also for each of the four sets of problems. The H average per cent for Good Explanations is 14 points or more larger than that of the L subjects for each set.

Proficiency. (a) The H and the L subjects revealed equal proficiency in work with the eight combinations, at least as far as Accuracy is concerned. (Table 116.) Too, they gave their answers quickly with about the same relative frequency; but the L group is charged with significantly fewer Long Pauses. The L group also has the larger average per cent for Correct Answers for three pairs of combinations, but the largest difference is 8 points, and has the smaller average per cent for Prompt Responses for three pairs by amounts never larger than 7 points. The H group has the larger average per cent for this category for the fourth pair by 8 points.

(b) The records of the two groups for Accuracy in computing answers for the four examples are about the same, though the H subjects much more frequently did their work steadily and smoothly. (Table 117.) None of the differences between the pairs of per cents for Correct Answers amounts to more than 4 points for any of the examples. For each example the H subjects have the larger per cent for Smooth Performances, the differences in their favor being 8 points in two, 12 points in one, and 15 points in one.

(c) Neither group surpassed the other in Accuracy in solving the twelve problems as a whole; but the H group has significantly the better record for High-Level Performances. (Table 118.) For none of the four sets of problems is the difference between the average per cents of the two groups for Correct Answers larger than 3 points; but in all four sets the H subjects have the larger average per cent for High-Level Performances, in three sets by significant margins.

B. The relative effectiveness of the three programs when taught for equal amounts of time.

In the study of this problem the relative effectiveness of the programs is to be determined both when the total amount of teaching time was relatively large and when it was relatively small. Whether large or small, the amounts of instruction offered in the three programs are not exactly equal, -- only approximately so.

Two sets of comparisons are required: (1) Cui. H vs Dien. H vs Con. H and (2) Cui. L vs Dien. L vs Con. L.

(4) Cui. H vs Dien. H vs Con. H

Conceptual maturity. (a) Perhaps the most convenient way to get a general assessment of the conceptual maturity of the three H groups with respect to their methods of thinking answers for the eight combinations is to compare their average per cents, rubric by rubric. The following summary is based upon the data in Table 113. The average per cents for Rubric C are included only for completeness and will be disregarded as usual in gross comparisons.

Rubrics	Average Per Cents		
	Cui. H	Dien. H	Con. H
A. Non-Attempts	11	10	2
B. Counting	30	28	26
C. Guessing, Rote Recall, Recitation of Tables, etc.	23	25	16
D. Rational Solutions	29	31	28
E. Uses of Meaningful Habituations	7	3	27

By reason of their substantially smaller average per cent for Rubrics A and B combined and their much larger average per cent for Rubric E, the Con. subjects can be said to have clearly outclassed each of the other two groups of subjects in conceptual maturity. There being only inconsequential differences between the

average per cents of the Cui. and the Dien. groups for the various rubrics, it is to be inferred that they were at about the same level of conceptual maturity.

Since the records of three pairs of groups are to be compared, differences between average per cents in the selected rubrics for the two combinations in each operation are reported in a manner that needs to be explained, especially since it will be used several times later on. The entries for the addition pair in the first row are read thus:

Rubrics A + B: The Con. group has a 6 - 9 point smaller average per cent than does either of the other groups, whose average per cents are equal or nearly so.*

Pairs of Combinations	**Rubrics**		
	A + B	**D**	**E**
Addition	**(Con.**)** **- -**	**Dien.*, Con.*** **- -**	**- -** **- -**
Subtraction	**(Cui.**)** **- -**	**- -** **- -**	**- -** **- -**
Multiplication	**Con.**** **Dien.***	**Con.**** **- -**	**Con.**** **- -**
Division	**Con.**** **- -**	**Con.**** **(Cui.*)**	**Con.**** **(Cui.*)**

Rubric D: The Con. and the Dien. groups, with about the same average per cents, both surpassed the Cui. group by significant margins.

Rubric E: The three groups have very similar records.**

---

[*] **The ()'s in (Con.**) indicate an insignificant difference; the **, that the difference holds in both comparisons; the --, that the average per cents of the Cui. and the Dien. groups are identical or nearly so.**

[**] **Differences amount to 0 or to no more than 5 points.**

These comparisons for pairs of combinations show that the evaluation of the three groups above with respect to conceptual maturity is almost certainly correct.

(b) The average per cents of the three groups for Satisfactory Explanations of the four computations are 46 for the Cui.subjects, 44 for the Dien. subjects, and 66 for the Con. subjects. (Table 114.) According to this evidence, the Con. group demonstrated by far the highest degree of conceptual maturity, with nothing to choose between the Cui. and the Con. groups.

In every one of the four separate examples the Con. per cent for Satisfactory Explanations is the largest of the three by differences never less than 13 points. The largest difference between the corresponding measures of the Con. and the Dien. groups is 5 points, and it occurs only once. The results of this example-by-example analysis underscore, if they do not increase, the margin by which the Con. subjects outmatched the subjects in the other two groups in the ability to account mathemetically for computations.

(c) The Con. group gave Good Explanations of their solutions of the twelve problems in 75% of their attempts; the Cui. group, in 39%; and the Dien. group, in 50%. (Table 115.)

The average per cents of the three H groups for Good Explanations for the four sets of problems are, in the order Cui., Dien., Con.:

Addition	47; 72; 81	Multiplication	40; 42; 77
Subtraction	33; 47; 74	Division	38; 41; 67

According to the comparisons of averages for the whole twelve problems, the ranking of the groups in conceptual maturity is: Con., Dien., and Cui.; and all differences between ranks are significant. According to the comparisons for the sets of problems, the Con. subjects clearly maintain their right to first place by having the highest average per cent for Good Explanations in every set, -- in three by significant margins and in the fourth by almost a significant margin; and the claim of the Dien. subjects to second place is warranted, but only because of significantly larger average per cents for two sets in comparison with the Cui. subjects.

Proficiency. (a) According to the average per cents in Table 116, the Con. subjects have a 12-point advantage over the Dien. subjects and a 7-point advantage over the Cui. subjects in the category Correct Answers; and they have an 8-point advantage over both groups of subjects in the category Non-Attempts (that is, they made 8% more attempts to give answers). In Accuracy, as a whole, then, the Con. group scored highest, and the scores of the Cui. and the Dien. groups indicate equality in this respect.

The comparative records of the three groups with pairs of combinations in the categories Correct Answers and Non-Attempts bear out this evaluation. The following summary includes only significant and nearly significant differences -- the latter indicated by ()'s -- between pairs of average per cents.

Pairs of Combinations	Correct Answers	Non-Attempts
Addition	- - - -	- - - -
Subtraction	(Cui.*) (Con.*) Dien.	- - - -
Multiplication	Con.** - -	(Con.**) - -
Division	Con.** - -	Con.** - -

To return to Table 116: the Con. group is credited with slightly more Prompt Responses to the eight combinations than the Dien. group and with significantly more such responses than the Cui. group. However, the Dien. subjects are charged with the smallest proportion of Long Pauses by substantial margins over the subjects in the other two groups. The corresponding data for the pairs of combinations are omitted since they contribute nothing of importance.

All in all, in proficiency in work with the number combinations the Con. group rates highest, and the Cui. and the Dien. groups are tied for second place.

(b) The Con. group was the most proficient of the three groups in computing answers for the four examples taken together, having the largest average per cents for Correct Answers and for Smooth Performances and the smallest average per cents for Non-Attempts and for Erratic Performances. (Table 117.) All except two of the eight differences for pairs of average per cents are significant, and the two exceptions are nearly so (9 percentage points in each case). The Cui. and the Dien. groups were equally accurate, but the Cui. subjects worked steadily and smoothly considerably more often than did the Dien. subjects.

The per cent of the Con. group for Correct Answers is substantially larger than those of the other two groups for each of the four examples, and the average per cent of the Cui. group is substantially larger than that of the Dien. group for one example only. The Dien. subjects were somewhat more accurate than the Cui. subjects in solving one example.

In per cent for Smooth Performances the Con. subjects surpassed the Cui. subjects by substantial margins in three examples and surpassed the Dien. subjects in all four. The Cui. subjects have a significant advantage in this category over the Dien. subjects in all four examples.

In sum, the Con. H group was much the more proficient of the three in computation, with little difference in this respect between the Cui. H and the Dien. H groups.

(c) For Correct Answers for the twelve problems the Con. average per cent is somewhat larger than that of the Dien. subjects and is substantially larger than that of the Cui. subjects. (Table 118.) For High-Level Performances the Con. average per cent is considerably larger than that of either of the other groups.

The summary below reports the significant and the near-significant differences between pairs of average per cents for the four sets of problems.

Problem Sets	Correct Answers	High-Level Performances
Addition	(Con.*) - -	Con.** Dien.*
Subtraction	(Con.*) - -	Con.** Cui.*
Multiplication	Con.* (Con.*) Dien.*	Con.** - -
Division	Con.* (Con.*) (Dien.*)	Con.** Cui.*

To conclude, the rank order for proficiency in problem solving is Con., Dien., Cui. The superiority of the Con. and the Dien. subjects over the Cui. subjects is large; that of the Con. subjects over the Dien. subjects, not so large.

(5) Cui. L vs Dien. L vs Con. L

Conceptual maturity. (a) As is shown in the following summary of average per cents from Table 113, the Con. subjects employed thought processes at the lowest levels of conceptual maturity (Rubrics A and B combined) significantly less often than did either of the other two groups of subjects in finding answers for the eight combinations. And they employed thought processes at the highest level (Rubric E) very considerably more often. By having a 6-point smaller average per cent for Rubrics A and B together and a 10-point larger average per cent for Rubric D, the Cui. subjects displayed a somewhat higher degree of conceptual maturity than did the Dien. subjects.

Rubrics	Average Per Cents Cui. L	 Dien. L	 Con. L
A. Non-Attempts	11	2	2
B. Counting	24	39	19
C. Guessing, Rote Recall, Recitation of Tables, etc.	29	28	31
D. Rational Solutions	33	23	28
E. Uses of Meaningful Habituations	1	5	19

The advantage of one or another of the groups for each pair of combinations is shown in the following summary.

Pairs of Combinations	Rubrics A + B	 D	 E
Addition	Cui.*, Con.* - -	Dien.*, Con.* - -	- - - -
Subtraction	(Cui.**) - -	Con.** - -	- - - -
Multiplication	Con.** - -	Dien.*, Cui.* - -	Con.** (Dien.*)
Division	Con.** (Dien.*)	Dien.* (Dien.*) Cui.*	Con.** - -

The fact that the Con. group has the smallest of the three average per cents for Rubric D for the multiplication and for the division pairs is accounted for by their relatively much larger average per cents for Rubric E, -- 39 and 32, respectively, to be compared with 10 and 2, the larger of the two average per cents of the other two groups. Hence, the ranking of the three groups in degree of conceptual maturity in terms of averages for the eight combinations as a whole is confirmed, with the Con. group in first place and the Cui. group in second.

(b) The average per cents of the three L groups for Satisfactory Explanations of the four computations are: Cui., 21; Dien., 45; Con., 68. (Table 114.)

The per cent of the Con. subjects for Satisfactory Explanations is the largest of the three for each of the four separate computations. In all eight comparisons with the Cui. and the Dien. subjects, the difference in favor of the Con. subjects never amounts to fewer than 19 points. And the per cent of the Dien. subjects for this category exceeds that of the Cui. subjects in each computation by at least 20 points.

In degree of conceptual maturity, as attested in this instance by relative ability to account mathematically for computations, the Con. group easily rates highest and the Dien. group next highest.

(c) The Dien. subjects gave Good Explanations of their solutions of the twelve problems, on the average, in 67% of their attempts, compared with the average per cents of 49 for the Cui. subjects and 48 for the Con. subjects.

In similar comparisons for the four sets of problems the Dien. subjects held to their highest rating for Good Explanations. The average per cents for this category, set by set, in the order Cui., Dien., and Con., are:

Addition 71; 89; 66 Multiplication 52; 64; 48
Subtraction 44; 66; 47 Division 28; 49; 33

By substantial margins, then, the Dien. group surpassed both the Cui. and the Con. groups in conceptual maturity, and the latter two groups are to be rated as at about the same level in this respect.

Proficiency. (a) The Con. and the Dien. groups both attempted answers for the eight combinations, on the average, in 97% of their opportunities, and the Cui. group in 84%; and the average per cents of the three L groups for Correct Answers are: Con., 83; Dien., 78; Cui., 62. (Table 116.)

For the pairs of combinations the Accuracy records of the groups, after the usual evaluation for significance and for near-significance, are:

Pairs of Combinations	Correct Answers	Non-Attempts
Addition	Dien.* Con.* Cui.	- - - -
Subtraction	Con.** Dien.*	Con.* Dien.*
Multiplication	Con.* (Con.*) Dien.*	(Dien.*) (Con.*) Cui.
Division	Dien.* Con.* Cui.	Dien.* Con.* Cui.

The relative inaccuracy of the Cui. subjects is observable in the data both for Correct Answers and for Non-Attempts. Yet, the Cui. subjects have the largest of the three average per cents for Prompt Responses to the eight combinations, by 30 points over the Dien. subjects and by 8 points over the Con. subjects (Table 116), -- evidence enough that measures of Rate are untrustworthy indices of proficiency except when Accuracy is at a high level. For this reason the data for Prompt Responses for the pairs of combinations are omitted.

In proficiency with the number combinations the Con. group outranks the Dien. group somewhat, but both these groups greatly outrank the Cui. group.

(b) On the average, the Con. subjects tried to solve the four examples in 97% of their opportunities and achieved Correct Answers in 79% of their attempts. (Table 117.) The corresponding per cents for the Dien. subjects are 87 and 63; for the Cui. subjects, 91 and 44. The Con. group therefore made the best of the three records for Accuracy, and the Dien. group has a much better record than has the Cui. group.

The per cents for Correct Answers and for Non-Attempts for the separate

computations support this relative standing of the three groups with respect to Accuracy.

Examples	**Correct Answers**	**Non-Attempts**
27 + 46	**Dien.* Con.*** **- -**	**- -** **- -**
92 - 45	**Con.**** **(Dien.*)**	**Con.**** **Dien.***
3 x 28	**Con.**** **Dien.***	**Con.*** **(Cui.*)**
84 ÷ 4	**Con.**** **Dien.***	**Con.**** **(Cui.*)**

The average per cents for the three L groups for Smooth Performances in the four computations are: Con., 78; Cui., 75; Dien., 62. In this category the Con. subjects surpassed the Dien. subjects in three examples by substantial margins and the Cui. subjects in two examples.

These proficiency data being of less importance than are those for Accuracy, the rating of the three groups in computational proficiency remains the same as reported above.

(c) The Con. subjects attempted solutions of the twelve problems, on the average, in 99% of their chances and were successful in 86% of their attempts. For the Dien. subjects the corresponding average per cents are 98 and 87; for the Cui. subjects, 90 and 67. (Table 118.)

Comparisons of the average per cents for these two categories for the four sets of problems resulted in the following ratings.

Problem Sets	Correct Answers	Non-Attempts
Addition	Dien.* Con.* Cui.	- - - -
Subtraction	Dien.* Con.* Cui.	Dien.* Con.* Cui.
Multiplication	Dien.* Con.* Cui.	(Con.*) - -
Division	Dien.* Con.* Cui.	Con.* (Dien.*) Cui.

For the category High-Level Performances for the twelve problems the Con. group has much the largest of the three average per cents, and those for the Cui. and the Dien. groups differ little.

All in all, the Con. and the Dien. L groups exhibited equal skill in solving the problems, and both these groups were much more skillful than was the Cui. group.

Summary, Section 1

Conceptual maturity

As a simple method of comparing groups with respect to relative conceptual maturity, Table 118A has been constructed, so numbered, somewhat illogically, in order to preserve the parallel numbering of tables in the Scottish and the English studies. In the Scottish study no such summary table was necessary because the relative effectiveness of only two programs was to be determined. There will be a similarly irregular table, Table 118B, to assemble the results of comparisons of groups for proficiency.

Intra-program comparisons. In only two of the nine comparisons in (1) to (3) did the H group significantly surpass the L group in conceptual maturity. In four others the H and the L groups were judged to be equal in this respect.

Table 118A. Time as a Factor. Relative Standing of Groups in Conceptual Maturity*

Groups Compared	Thought Processes with Combinations	Ability to Explain Computations	Ability to Explain Solutions of Problems
(1) Cui. H vs Cui. L	Equal	H	L
(2) Dien. H vs Dien. L	(L)	Equal	L
(3) Con. H vs Con. L	Equal	Equal	H
(4) Cui. H vs Dien. H vs Con. H	1. Con. 2. Cui. and Dien. equal	1. Con. 2. Cui. and Dien. equal	1. Con. 2. Dien. 3. Cui.
(5) Cui. L vs Dien. L vs Con. L	1. Con. 2. Cui. 3. Dien.	1. Con. 2. Dien. 3. Cui.	1. Dien. 2. Cui. and Con. equal

* In (4) and (5) no distinction has been made between differences that are significant and those that are slightly less than significant.

In the three remaining comparisons the advantage lies with the L group, twice significantly so. In the light of these facts, the larger amounts of instruction offered the H groups, never less than 52% of the total time given their paired L groups, failed rather generally to lead to correspondingly greater progress toward high levels of arithmetical abstractions; and time, as here defined, was therefore scarcely an important factor in influencing the effectiveness of any of the programs in promoting conceptual maturity.

Yet, this inference may be premature in that it does not take into account differences in factors other than time. In (1), the L subjects were brighter and more successful in school learning than were the H subjects. The differences in their favor amount to about 0.25 of a rating in each of the following: brightness, over-all achievement in arithmetic, and achievement in reading. Indeed, the greater learning ability of the L subjects may be the reason (a) why they were given less instruction and (b) why in three instances they equalled or surpassed the H subjects in conceptual maturity. If this is so, the results of the comparisons in (1) may have had little to do with the evaluation of time as a factor in the selected Cui. schools.

In (2) the H subjects were estimated to be brighter than the L subjects (0.19 of a rating) and to have achieved better in reading (0.09 of a rating) but less well in arithmetic (0.13 of a rating). Their two advantages in these ratings, coupled with their advantage of 80% in amount of instruction, should seemingly have enabled the H subjects to excel in conceptual maturity; yet, they did not. In the data collected there is nothing to explain why this was so, though one may guess that the L subjects were better taught. Whatever the reason or the reasons, in (2) time was definitely not a factor in determining the effectiveness of the Dien. program.

In (3) the H group was estimated to be much brighter than the L group (0.32 of a rating) and to have been much better achievers both in arithmetic (0.31 of a rating) and in reading (0.38 of a rating). Add to their evident superiority in learning ability the fact that the H subjects had had 66% more teaching time, and one is at a loss to understand why only in explaining solutions of problems did they reveal a degree of conceptual maturity greater than that of the L subjects. In (3), then, time can be minimized as a factor that conditioned the effectiveness of the Con. program.

Cross-comparisons of programs. Unlike the purpose in (1), (2), and (3), the purpose in (4) and (5) is to assess the relative effectiveness of the three programs, -- in (4), when the three groups had had relatively large amounts of instruction; in (5), when the three groups had had considerably less instruction in arithmetic.

(4) According to Table 118A the Con. H subjects revealed greater conceptual maturity than did either the Cui. H or the Dien. H subjects in dealing with each of the three types of arithmetical task, and their advantages in all comparisons were found to be significant. Except for the substantial superiority of the Dien. subjects in ability to explain solutions of problems, the Cui. and the Dien. subjects seem to have attained about the same level of conceptual maturity. The inference from these facts may be stated something like this: Granted a large amount of instructional time, the same for all three programs, the Con. program was by far the most effective in promoting conceptual maturity, and the Cui. and the Dien. programs were each about as effective as the other.

This inference must be tentative pending an analysis of the relative equality *of the H groups in various respects. The basic table,* Table 112, being as remote as it is, certain of the data therein will be repeated.

Weighted Means	Cui. H	Dien. H	Con. H
Time	720	764	736
Age	7 - 10	8 - 1	7 - 11
Ratings for			
Brightness	2.32	2.35	1.98
Arithmetic Achievement	2.35	2.38	1.94
Reading Achievement	2.44	2.65	2.02

The large advantages of the Con. group in brightness and in school learning raise serious doubts concerning the interpretation given above of the results of the comparisons in (4). The subjects in the Con. schools may have progressed further in the movement toward high levels of arithmetic abstractions, not because of especially commendable and possibly unique features of the Con. program as such, but because they were better learners. The latter explanation seems preferable to the former.

(5) As is shown in Table 118A, the Con. L group has first place in ranking for conceptual maturity exhibited in two of the three arithmetical tasks and is tied with the Cui. L group for second place in one. The Dien. L group has the highest rating in one, next to the highest in one, and the lowest in one. The Cui. L group was unable in any of the three tasks to get a position higher than second. On the whole, the Con. program seems to have been the most effective of the three, but by no large margin over the Dien. program. There is little question, however, about assigning the Cui. program the lowest place.

Again, this assessment must be examined in the light of data from Table 112.

Weighted Means	Cui. L	Dien. L	Con. L
Time	400	424	483
Age	7 - 6	8 - 2	8 - 0
Ratings for			
Brightness	2.08	2.54	2.30
Arithmetic Achievement	2.11	2.25	2.29
Reading Achievement	2.19	2.56	2.40

The Con. subjects had had 14% more instruction than the Dien. subjects, and in addition they were estimated to be brighter and somewhat better readers. On the other hand, in this investigation differences in teaching time as small as 14% have not been found decisive in determining relative success in learning; and the difference in brightness in favor of the Con. subjects, 0.24 of a rating, may not have been large enough to account to any great extent for their better record. Nevertheless, even if the conclusions reached in this exercise in probabilities are acceptable, confidence that the Con. program was more effective than the Dien. program is somewhat shaken.

The Cui. program has been rated the lowest of the three programs in effectiveness. Yet, the subjects in the Cui. L group were much younger than those in the Dien. and the Con. groups and had had the least amount of teaching time, -- 17% less than the Con. subjects. On the other hand, their advantages in brightness and demonstrated ability in school learning should, it would seem, have compensated for their handicaps as just noted. In line with this reasoning, the rating of the Cui. program for effectiveness is not changed.

Proficiency

The reader is reminded yet again that proficiency in this investigation means proficiency only in dealing with the interview items in each arithmetical task.

The data for Rate or Quality (or Level) of Performance being misleading, as has been shown several times, those for Accuracy, and particularly those for Correct Answers, will be used in comparing groups for proficiency. See Table 118B.

In (1) to (3) the H group was more proficient than the L group in three comparisons (the differences in two are not significant), was less proficient in two, and equalled the L group in proficiency in four. Evidently, advantages

Table 118B. Tima as a Factor. Relative Standing of Groups in Proficiency*

Groups Compared	Number Combinations	Computation	Problem Solving
(1) Cui. H vs Cui. L	H	(H)	(H)
(2) Dien. H vs Dien. L	L	L	Equal
(3) Con. H vs Con. L	Equal	Equal	Equal
(4) Cui. H vs Dien. H vs Con. H	1. Con. 2. Cui. and Dien. equal	1. Con. 2. Cui. and Dien. equal	1. Dien. and Con. equal 3. Cui.
(5) Cui. L vs Dien. L vs Con. L	1. Con. 2. Dien. 3. Cui.	1. Con. 2. Dien. 3. Cui.	1. Dien. and Con. equal 3. Cui.

* As in Table 118A, in (4) and (5) no distinction is made between differences that are significant and those that are slightly less than significant.

in time, by themselves, were no more decisive in determining relative proficiency than they proved to be in determining relative degree of conceptual maturity.

Of the comparisons of the H groups in the three programs with regard to proficiency, (4), two favor the Con. group over the other two groups and, on the whole, favor neither the Cui. nor the Dien. group in comparison with the other.

In (5), in which the groups had had relatively small amounts of instruction, the over-all ranking of the three programs for producing proficiency is: Con., Dien., Cui.

All these evaluations are of course subject to the effects of differences in factors other than time, as was true for the evaluations for conceptual maturity.

2. Brightness as a Factor

The differences in the size of samples in the H groups and in the L groups, shown in the first row of Table 119, are of some interest. The H groups, rated 1 in brightness, represent 34% of the total number of subjects in the Cui. program, 24% of those in the Dien. program, and 35% of those in the Con. program. And the corresponding per cents for the L groups, rated 3 and 4 in brightness, are 35, 48, and 39. Either the total samples were dissimilar in this factor, or the scheme for estimating brightness was not uniformly applied to the program samples.

The disparities in weighted means for total instructional time given the six groups are not large enough to have affected materially their relative success in learning. The largest difference is that between these measures for the Cui. H and the Con. L groups, and it amounts to about 12% of the weighted mean of the group that had been taught for the shorter total time. The next largest difference, computed in the same way, is 7%.

All six groups have average ages within a range of three months.

Finally, the H groups in all program samples, have, as expected, much higher ratings for achievement both in arithmetic and in reading.

A. The influence of wide differences in brightness upon the effectiveness of each program.

(1) Cui. H vs Cui. L

Conceptual maturity. (a) The H subjects attempted answers for the eight combinations far more frequently than did the L subjects and counted to get answers far less frequently (Rubrics A + B, table 120). They also employed Rational Solutions more than twice as often as did the L subjects, and they have an average per cent for Uses of Meaningful Habituations that is almost significantly

Table 119. Brightness as a Factor. Composition of the H and the L Groups in Each Program

Bases of Comparison	Cuisinaire		Dienes		Conventional	
	H	L	H	L	H	L
Number of Subjects	87	95	85	171	108	121
Per Cents by Sex						
Male	43	52	53	51	58	55
Female	57	48	47	49	42	45
Weighted Means for						
Total Instructional Time	561	604	603	589	629	582
Age	7 yr. 10 mo.	8 yr. 0 mo.	8 yr. 0 mo.	8 yr. 1 mo.	8 yr. 0 mo.	7 yr. 11 mo.
Estimated Ratings:						
General Achievement in Arithmetic	1.34	3.17	1.20	3.04	1.19	3.03
Achievement in Reading	1.35	3.28	1.32	3.21	1.26	3.15

larger than that of the L subjects.

Next the records of the H and the L groups may be examined for the pairs of combinations. According to the summary below, for each of the four pairs the H subjects used Rubrics A and B together significantly less often than did the L subjects. For each pair, too, the H group has the larger average per cent for Rational Solutions (Rubric D), in all by substantial margins. For each pair the average per cent of the H group for Uses of Meaningful Habituations is the larger, for two by significant differences, for two by nearly significant

Table 120. The Number Combinations. Thought Processes with the Eight Combinations; Averages in Per Cents

Rubrics and Categories	Cuisinaire		Dienes		Conventional	
	H	L	H	L	H	L
A. Non-Attempts	5	15	5	17	1	4
B. Counting, All Forms	12	40	21	39	17	27
C. Guessing	2	6	2	7	2	9
Rote Recall	7	5	10	3	2	1
Recitation of Tables	10	7	7	7	10	12
Mechanical Uses: Reverses and Inverses	5	6	5	2	2	6
Sub-Totals	24	24	24	19	16	28
D. Rational Uses:						
Reverses	1	0	0	0	2	1
Inverses	3	0	3	1	2	2
Inverses, Plus	4	2	4	2	6	5
Solutions, Other Forms	38	16	34	18	23	17
Sub-Totals	46	18	41	21	33	25
E. Uses of Meaningful Habituations	11	2	8	3	33	14

differences (actually 7 and 9 points for the multiplication and the division pairs).

The evidence as a whole is interpreted to indicate that the H subjects were more advanced conceptually than were the L subjects in ways of thinking answers for the number combinations.

Pairs of Combinations	Rubrics		
	A + B	D	E
Addition	H	H	H
Subtraction	H	H	H
Multiplication	H	H	(H)
Division	H	H	(H)

(b) The H average per cent for Satisfactory Explanations of the four computations is more than double the size of the L average per cent (Table 121), and the H per cents for this category are the larger by 23 points or more for each of the examples.

(c) The H group displayed almost as much superiority over the L group as in (b), in verbalizing successfully solutions of the twelve problems. (Table 122.) The average per cents of the two groups for Good Explanations for the sets of problems are:

Addition	H, 93; L, 84	Multiplication	H, 86; L, 63
Subtraction	H, 87; L, 68	Division	H, 84; L, 54

Proficiency. (a) On the average, the L group attempted answers in 82% of their chances with the eight number combinations; the H group, in 96%; and the L group was considerably the less successful of the two in furnishing Correct Answers and in supplying answers promptly, whether correct or incorrect. (Table 123.)

For all four pairs of combinations, the H average per cent is the larger of the two for Correct Answers by at least 13 points and for Prompt Responses by at least 30 points.

(b) In computational skill as exhibited in solving the four examples combined, the H group considerably outclassed the L group, both in Accuracy and

Table 121. Computation. Quality of Explanations of Computations, Attempts Only; Averages in Per Cents

Classification of Explanations*	Cuisinaire		Dienes		Conventional	
	H	L	H	L	H	L
Satisfactory						
Words Only	12	4	28	15	29	13
Altered Examples	36	17	25	15	38	29
Other	0	0	2	3	0	0
Unsatisfactory						
Words Only	32	42	26	40	15	30
Altered Examples	19	25	9	8	17	26
Other	0	11	9	18	0	2

* The category Imagery is omitted

Table 122. Problem Solving. Reasons Given for Solutions, Attempts Only; Averages in Per Cents

Quality of Explanation*	Cuisinaire		Dienes		Conventional	
	H	L	H	L	H	L
Good	60	34	64	44	68	40
Unsatisfactory	16	24	7	26	5	20
Indeterminate	2	5	3	3	5	9

* The category Fair is omitted

Table 123. The Number Combinations. Proficiency in Giving Answers: Accuracy and Rate; Averages in Per Cents

Nature of the Measures	Cuisinaire		Dienes		Conventional	
	H	L	H	L	H	L
Accuracy						
Correct Answers	85	61	84	59	90	70
Non-Attempts	4	18	6	17	1	7
Rate, Attempts Only						
Prompt Responses	48	13	51	22	59	29
Long Pauses	12	40	14	38	10	29

Table 124. Computation. Proficiency: Accuracy and Quality of Performance; Averages in Per Cents

Types of Measure	Cuisinaire		Dienes		Conventional	
	H	L	H	L	H	L
Accuracy						
Correct Answers	72	45	70	35	87	63
Non-Attempts	5	21	11	36	1	5
Quality of Performance, Attempts Only						
Smooth	86	66	72	40	90	70
Erratic	14	34	28	60	10	30

Table 125. Problem Solving. Proficiency: Accuracy and Performance Levels; Averages in Per Cents

Types of Measure	Cuisinaire		Dienes		Conventional	
	H	L	H	L	H	L
Accuracy						
Correct Answers	87	67	89	72	93	78
Non-Attempts	1	10	2	8	1	3
Level of Performance, Attempts Only						
High	52	26	50	27	74	49
Low	13	49	9	37	8	23

in Quality of Performance. (Table 124.) Too, the brighter subjects have the smaller per cent for Non-Attempts for each of the examples (three differences are significant) and the larger per cents for Correct Answers and for Smooth Performances for all four, by a significant difference in each comparison.

(c) In solving the twelve problems the H record is much better than the L record in both measures of Accuracy and in High-Level Performances. (Table 125.) The H average per cents are consistently larger than the L average per cents for Correct Answers and for High-Level Performances, by significant amounts for all four sets of problems; and they are significantly smaller for Non-Attempts for two sets.

(2) Dien. H vs Dien. L

Conceptual maturity. (a) As the average per cents in Table 120 show, the L group failed to try answers for the eight combinations and counted to secure answers to a far greater extent than did the H group. Neither group made use of

Meaningful Habituations at all commonly, but the H subjects have a slight advantage in Rubric E. Their advantage in Rubric D is, however, substantial.

Following is a summary of the significant and near-significant differences between the groups in relative use of Rubrics A + B, D, and E with the pairs of combinations. Rather generally, the H subjects have the advantage, and therefore they are to be regarded as superior to the L subjects in conceptual maturity of methods employed to think answers to the combinations.

Pairs of Combinations	Rubrics		
	A + B	D	E
Addition	H	H	- -
Subtraction	H	H	- -
Multiplication	H	- -	H
Division	H	H	(H)

(b) The H subjects surpassed the L subjects in average per cents for Satisfactory Explanations of the four computations as a whole, and they were significantly more successful whether they used Words Only or Altered Examples. (Table 121.) The H per cent for Satisfactory Explanations exceeds the L per cent for each of the separate computations by amounts equal to 23 points or more.

(c) The L group was much less able than was the H group to offer Good Explanations of the solutions of the twelve problems combined (Table 122) and of the solutions in each of the four sets of problems. In the last comparisons, all differences between average per cents are significant.

Proficiency. (a) In work with the eight combinations the H subjects attempted answers considerably oftener than did the L subjects and were much more able to give Correct Answers (Table 123), at the same time that they responded promptly substantially more frequently.

Comparisons of the two groups for the pairs of combinations also show the greater proficiency of the H subjects, who have the larger average per cents both for Correct Answers and for Prompt Responses by wide margins for each pair.

(b) According to Table 124, the brighter children were clearly superior to those less bright in all comparisons for computational proficiency revealed in work with the four examples, as they were also in solving each of the examples. For each example the per cent of the H group for Correct Answers is at least 30 points larger than that of the L group, and for Prompt Responses, at least 27 points the larger.

(c) The H subjects also exhibited much the greater skill in solving the twelve problems. (Table 125.) Their average per cents for Correct Answers are consistently larger than those of the L subjects for the four sets of problems, significantly so for three sets; and their average per cents for High-Level Performances are the larger for each set, always by a margin of at least 20 points.

(3) Con. H vs Con. L

Conceptual maturity. (a) As usual, the H subjects thought answers for the eight combinations more commonly than did the L subjects, in ways rated toward the upper end of the conceptual scale. Besides counting substantially less often, the H subjects used Rational Solutions somewhat more frequently and Meaningful Habituations considerably more frequently. (Table 120.)

The significant and the near-significant differences between the average per cents of the H and L groups for the pairs of combinations are as below. As a whole, the comparisons are in favor of the H subjects and, together with the findings in comparisons for average per cents for the twelve combinations, they warrant the conclusion that the H subjects had reached a higher level of conceptual maturity than had the L subjects, in thought processes used to obtain answers for the combinations.

Pairs of Combinations	Rubrics A + B	D	E
Addition	H	H	(H)
Subtraction	- -	L	- -
Multiplication	H	- -	H
Division	(L)	- -	H

(b) The average per cents for Satisfactory Explanations of the four computations together are 67 for the H group and 42 for the L group (Table 121), and the differences between the per cents of the two groups for Satisfactory Explanations favor the H group in each of the computations and are never smaller in amount than 20 points.

(c) According to Table 122, the H subjects demonstrated far greater ability than did the L subjects, in accounting for solutions of the twelve problems, earning credit for Good Explanations almost twice as often. The relative superiority of the H subjects is so marked because their average per cent for Good Explanations is at least 20 points the larger for each of the four sets of problems.

Proficiency. (a) In terms both of Accuracy and of Rate, the H group made a much better record for proficiency in dealing with the eight combinations than did the L group. (Table 123.) The H average per cent for Correct Answers is larger than the L average per cent by 6 points only for the addition pair of combinations but by 16 or more points for the other three pairs; and the H average per cent for Prompt Responses is the larger for all four pairs, in each by at least 24 points.

(b) The H subjects were much the more skillful in computing answers for the four examples combined. For all four categories in Table 124, the comparisons favor the H group, by substantial margins in three. The same group exhibited much the greater proficiency in solving each example by consistently and considerably outclassing the L groups in comparisons for Correct Answers and for Smooth Performances.

(c) In much the same fashion the H subjects outmatched the L subjects in skill in solving the twelve problems. Comparisons of the two groups for the four categories in Table 125 all favor the H group, and the margins of difference in three are much larger than is required to designate them as significant. The average per cent of the H group considerably exceeds that of the L group for Correct Answers for three of the four sets of problems, and for High-Level Performances for all four sets.

B. The relative effectiveness of the three programs when taught to children rated 1 in brightness and when taught to children rated 3 and 4 in this factor.

(4) Cui. H vs Dien. H vs Con. H

Conceptual maturity. (a) The data below, taken from Table 120, will be useful in comparing the three H groups with respect to conceptual maturity in ways of thinking answers for the eight combinations taken together.

The data for Rubric C, which will be disregarded as usual in comparisons of this kind, are presented only to complete the records. And the data for Rubric D are of little value as indices of conceptual maturity because, as has been noted, average per cents for this rubric tend to decrease in size when, as here for the Con. group, the average per cents for Rubric E are reasonably large.

Rubrics	Average Per Cents		
	Cui. H	Dien. H	Con. H
A. Non-Attempts	5	5	1
B. Counting	12	21	17
C. Guessing, Rote Recall, Recitation of Tables, etc.	24	24	16
D. Rational Solutions	46	41	33
E. Uses of Meaningful Habituations	11	8	33

This means that in (4), the evaluation of the three programs with respect to effectiveness in promoting conceptual maturity must be determined from the data for Rubrics A, B, and E.

The average per cents for Rubrics A and B combined are: Cui., 17; Dien., 26; Con., 18. The Cui. and the Con. groups are charged with the use of thought processes low in the scale of conceptual maturity to the same extent; but both groups employed such thought processes somewhat less commonly than did the Dien. group.

The Con. group made use of Meaningful Habituations three times as often as did the Cui. group and four times as often as did the Dien. group, a fact that should be of considerable importance here in view of the fact that this category or rubric is rated highest in the scale of conceptual maturity.

In terms of these comparisons of average per cents the relative standing of the three groups in respect to conceptual maturity of thought processes is: Con., Cui., Dien. However, this ranking is based upon small amounts of difference, except for Rubric E.

As a matter of fact, this ranking proves to be highly questionable when one examines the results obtained by comparing average per cents for pairs of combinations, summarized below in the usual manner.

Pairs of Combinations	Rubrics A + B	D	E
Addition	Cui.** - -	- - - -	(Cui.**) - -
Subtraction	Cui.* Con.*	Cui.** Dien.*	(Cui.*) - -
Multiplication	(Cui.*) - -	Cui.** Dien.*	Con.** Dien.*
Division	Con.** - -	Cui.* Dien.* Con.	Con.* - -

In the comparisons above the Cui. subjects have significant advantages in eight and nearly significant advantages in four others; and the Dien. and the Con. subjects have significant advantages in four and seven, respectively. Moreover, four of the seven advantages of the Con. subjects were earned in Rubric E for the multiplication and division pairs which they may well have studied more thoroughly than had the subjects in the other two groups.

In the light of all the facts, it is probably advisable to rate the three H groups as about equally matched in the level of conceptual maturity they displayed in their ways of thinking answers for the number combinations.

(b) For Satisfactory Explanations of the four computations the average per cents of the three H groups, taken from Table 121, are: Cui., 48; Dien., 55; Con., 67.

The per cent of the Con. subjects for Satisfactory Explanations is significantly larger than that of the subjects in each of the other groups for three computations and is almost significant for the fourth. The only other significant differences are two in number, both in favor of the Dien. group over the Cui. group.

The results of these example-by-example comparisons agree with those of the comparisons of averages, and there is little doubt that first place in conceptual maturity belongs to the Con. group and second place, to the Dien. group.

(c) The average per cents of the three H groups for Good Explanations of solutions of the twelve problems are: Cui., 60; Dien., 64; Con., 68. (Table 122.) Accordingly, in terms of averages the Con. subjects exhibited somewhat more ability than did the Cui. subjects in accounting for problem solutions and little, if any more, ability than did the Dien. subjects, who in turn were little, if any more, successful than were the Cui. subjects.

Comparisons of the average per cents of the three groups for Good Explanations for the sets of problems resulted as follows:

Addition set: Dien., one significant and one nearly significant advantage.

Subtraction set: Dien. and Con., one significant advantage each.

Multiplication set: Con., one significant and one nearly significant advantage; Cui., one nearly significant advantage.

Division set: Con., one nearly significant advantage.

On the whole, the evidence seems to mean that the Con. group ranks highest in conceptual maturity, but by a slight margin in comparison with the Dien. group and a somewhat larger margin in comparison with the Cui. group; and that the Dien. group ranks second by a small margin in comparison with the Cui. group.

Proficiency. (a) According to Table 123, the Con. group has percentage advantages of 6 points over the Dien. group and of 5 points over the Cui. group in Correct Answers for the eight combinations, negligible advantages over both groups in Non-Attempts, and advantages of 8 and 11 points over the Dien. and the Cui. groups, respectively, in Prompt Responses.

For the pairs of combinations, significant and near-significant differences between pairs of average per cents for the two categories under Accuracy are as below:

Pairs of Combinations	Correct Answers	Non-Attempts
Addition	- - - -	- - - -
Subtraction	(Con.*) - -	- - - -
Multiplication	(Con.*) - -	- - - -
Division	Con.* (Con.*) (Cui.*)	Con.* (Con.*) - -

While the Con. group has far the most differences in its favor, four of them occur in comparisons for the division pair, on which, as has been previously noted, the Con. group had apparently had more instruction than the other groups.

The Con. subjects have the largest average per cent for Prompt Responses for three pairs of combinations, and the differences in their favor are significant in five of the eight comparisons. The differences between the average per cents of the Cui. and the Dien. subjects in this category are never larger than 6 points.

Since the three average per cents for Correct Answers are 84 or larger, comparisons for Rate should probably be included in evaluating the three groups for proficiency in supplying answers for the number combinations. If this is done, the Con. group gets the highest rating in proficiency, and the Cui. and the Dien. groups are equal contenders for second place.

(b) The average per cent of the Con. group for Correct Answers for the four examples exceeds that of the Dien. group by 17 points and that of the Cui. group by 15 points. The three average per cents for Non-Attempts are small and differ little, and for Smooth Performances they are 86, 72, and 90 for the Cui., the Dien., and the Con. groups, respectively.

Comparisons of per cents for Accuracy of computation in the four examples yielded the following significant and near-significant differences.

Examples	Correct Answers	Non-Attempts
27 + 46	Con.** - -	- - - -
92 - 45	Con.** - -	(Cui.*) (Con.*) Dien.
3 x 28	Con.* (Con.*) - -	Cui.* Con.* Dien.
84 ÷ 4	Con.* (Con.*) (Cui.*)	Con.** - -

In four comparisons of per cents for Smooth Performances, the Cui. group has significant differences in its favor; in five, the Con. group has such advantages.

The results of the example-by-example comparisons support those of the comparisons of averages in assigning first place to the Con. group in computational skill. The claims of the Cui. and the Dien. groups to second place are of about equal merit.

(c) The Con. group differs little from the Dien. group in comparisons for frequency of Correct Answers and of Non-Attempts in solving the twelve problems; and, as a matter of fact, it was but slightly superior to the Cui. group, having a percentage advantage of only 6 points over the latter group in the category, Correct Answers. In relative frequency of High-Level Performances, however, the Con. group substantially surpassed both the other groups. (Table 125.)

For none of the four sets of problems is the average per cent of any group for Correct Answers significantly larger than are those of the other two groups. The largest difference is 7 points, and it occurs only twice. In all four sets, however, the Con. subjects, by considerable margins, surpass the Dien. and the Cui. subjects in relative frequency of High-Level Performances.

If the results of the comparisons for High-Level Performances are minimized, as they should be, the inference to be drawn is that the three groups were a match for each other in skill in problem solving.

(5) Cui. L vs Dien. L vs Con. L

Conceptual maturity. (a) The summary below of the average per cents of the three L groups for use of the various rubrics of processes in thinking answers for the eight combinations is taken from Table 120.

Rubrics	Average Per Cents		
	Cui. L	Dien. L	Con. L
A. Non-Attempts	15	17	4
B. Counting	40	39	27
C. Guessing, Rote Recall, Recitation of Tables, etc.	24	19	28
D. Rational Solutions	18	21	25
E. Uses of Meaningful Habituations	2	3	14

The Con. subjects employed thought processes of the two types rated lowest in meaningful abstractness (Rubrics A and B combined) in 31% of their attempts to obtain answers for the eight combinations; the Cui. subjects, in 55%, and the Dien. subjects, in 56%. In use of Rational Solutions the Con. group has a percentage advantage of 7 points over the Cui. group and of but 4 points over the Dien. group. While the Con. group made use of Meaningful Habituations in only 14% of its attempts, this average per cent greatly exceeds the average per cents of the two other L groups.

The results of comparisons of the three L groups in work with the pairs of combinations are presented below in the usual manner. Only significant and near-significant differences between average per cents are reported.

Pairs of Combinations	Rubrics		
	A + B	D	E
Addition	Con.** - -	Con.** - -	- - - -
Subtraction	Con.** - -	Con.** - -	- - - -
Multiplication	Con.** - -	Dien.* (Dien.*) (Cui.*)	Con.** - -
Division	Con.** - -	(Cui.**) - -	Con.** - -

To sum up, the Con. group displayed a stage of conceptual maturity much more advanced than that of either the Cui. or the Dien. group in ways of thinking answers for the combinations; and the Cui. group seems to have been at a slightly higher stage than the Dien. group.

(b) The average per cent of the Cui. group for Satisfactory Explanations of the four computations is 21, compared with 33% for the Dien. group and 42% for the Con. group. (Table 121.)

The Con. group has the largest of the three per cents for Satisfactory Explanations for each of the four computations. The differences in their favor are significant in three examples in comparison with the Cui. group and are significant in one example and nearly significant in another in comparison with the Dien. group. The last-named group has substantially larger per cents than the Cui. group in two examples.

The results of these last comparisons increase considerably the probability that the Con. group ranks highest in ability to account mathematically for computations. Second place goes to the Dien. group, though its advantage over the

Cui. group is not large.

(c) According to Table 122, the Dien. group surpassed the Con. group little if at all and surpassed the Cui. group substantially in ability to make clear the reasons for solutions of the twelve problems.

The relative success of the three groups in explaining solutions of sets of problems is shown by the following summary.

Addition set: Dien., two significant advantages; Con., one nearly significant advantage.

Subtraction set: Dien., one significant and one nearly significant advantage.

Multiplication set: no differences, significant or nearly significant.

Division set: Dien. and Con., one significant advantage each.

In sum, the Cui. group clearly rates lowest in the form of conceptual maturity here in question, and the Dien. group should probably be rated first, especially in view of the results of the set-by-set comparisons.

Proficiency. (a) The Con. group attempted answers to the eight combinations substantially oftener than did either of the other groups; has an 11-point advantage over the Dien. group and a 9-point advantage over the Con. group in average per cents for Correct Answers, and has advantages of 7 and 16 points over two groups in average per cents for Prompt Responses. (Table 123.)

Significant and near-significant differences in Accuracy with the pairs of combinations are:

Pairs of Combinations	**Correct Answers**	**Non-Attempts**
Addition	**- -** **- -**	**- -** **- -**
Subtraction	**- -**	**Con.*** **(Cui.*)**
Multiplication	**Con.**** **- -**	**Con.**** **- -**
Division	**Con.**** **(Cui.*)**	**Con.**** **- -**

It will be noted that eight of the nine advantages of the Con. group occur in connection with the pairs of multiplication and division combinations.

The Dien. and the Con. groups have average per cents of equal size for Prompt Responses for two pairs; but the average per cent of the Con. group is considerably the larger for each of the other two pairs. The record of the Cui. group is clearly the poorest of the three in this category.

Despite the possible implication of the statement made above with respect to the two special areas in which the Con. group displayed its greater amount of superiority in Accuracy, the order of proficiency in dealing with the number combinations is: Con., Dien., Cui.

(b) On the average, the Con. group attempted to solve the four examples in 95% of their opportunities; the Cui. group, in 79%; the Dien. group, in 64%. Further, the first-named group has the largest of the three average per cents for Correct Answers by dependable margins, and the average per cent of the Cui. group in this category exceeds that of the Dien. group by 10 points. The Con. subjects also surpassed the Dien. subjects, and the Cui. subjects surpassed the Dien. subjects, in comparisons for Smooth Performances. All the differences noted are in excess of 20 points. (Table 124.)

The significant and near-significant differences among the three L groups in Accuracy of computation for each of the four examples are summarized below.

Examples	Correct Answers	Non-Attempts
27 + 46	Con.** - -	Con.* (Cui.*)
92 - 45	Con.** - -	Con.** (Cui.*)
3 x 28	Con.** Cui.*	Con.** Cui.*
84 ÷ 4	Con.** Cui.*	Con.** Cui.*

For Smooth Performances the Dien. subjects made much the poorest of the three records in all four examples, and the record of the Con. subjects is significantly better than that of the Cui. subjects in one example and nearly so in a second.

The order of the three L groups in computational skill is: Con., Cui., Dien.

(c) As calculated from the average per cents in Table 125, the Con. subjects have differences in their favor in solving the twelve problems as follows:

Correct Answers: 11 points over the Cui. subjects; 6 points over the Dien. subjects.

Non-Attempts: 7 and 5 points over the Cui. and the Dien. subjects, respectively.

High-Level Performances: 23 and 22 points over the Cui. and the Dien. subjects, respectively.

For each of the four sets of problems the Con. average per cent for Correct Answers is the largest of the three, but their advantages are significant for only one set in comparisons with the Dien. subjects and for two sets in comparisons with the Cui. subjects. For none of the sets is the average per cent of the Dien. group for this category larger than that of the Cui. group. The Con. subjects far surpassed the Dien. and the Cui. subjects in average per cents for High-Level Performances in solving each of the four sets.

The Con. group seems to have been the most skillful in problem solving, followed closely by the Dien. group.

Summary, Section 2

Conceptual maturity

Table 125A summarizes the relative rankings of groups with respect to the level of conceptual maturity they exhibited in ways of thinking answers for number combinations, in successful attempts to account mathematically for computations in abstract examples, and in acceptable explanations of solutions of one-step problems.

Table 125A. Brightness as a Factor. Relative Standing of Groups in Conceptual Maturity*

Groups Compared	Thought Processes with Combinations	Ability to Explain Computations	Ability to Explain Solutions of Problems
(1) Cui. H vs Cui. L	H	H	H
(2) Dien. H vs Dien. L	H	H	H
(3) Con. H vs Con. L	H	H	H
(4) Cui. H vs Dien. H vs Con. H	Equal	1. Con. 2. Dien. 3. Cui.	1. Con. 2. Dien. 3. Cui.
(5) Cui. L vs Dien. L vs Con. L	1. Con. 2. Cui. 3. Dien.	1. Con. 2. Dien. 3. Cui.	1. Dien. 2. Con. 3. Cui.

* In (4) and (5) no distinction has been made between differences that are significant and those that are slightly less than significant.

Intra-program comparisons. Here interest centers upon the findings in (1) to (3) in Table 125A, in which children rated 1 in brightness are compared, program by program, with children rated 3 and 4 in this factor. It is hardly astonishing that in all nine comparisons the H groups excelled. That is to say, the Cui. program was more effective, as were the Dien. and the Con. programs, in promoting conceptual maturity when taught to children estimated to be very bright than when taught to children estimated to be much less bright.

The consistently better records of the H groups is not to be explained away as being the result of advantages they possessed in relative amounts of instruction. According to Table 119, the Cui. H group had had about 7% less teaching time than the Cui. L group; the Dien. H group, about 2% more than the Dien. L group; the Con. H group, about 8% more than the Con. L group. None of these differences is large enough to have affected learning to any appreciable extent.

Of course each H group was rated much higher than its paired L group for achievement both in arithmetic and in reading. On the other hand, these higher ratings in school achievement are typically the accompaniments of superior brightness, if indeed they are not the natural consequences thereof.

In conclusion, the consistent findings in (1), (2), and (3) indicate that brightness was an important factor in determining the effectiveness of each of the three programs.

Cross-comparisons of programs. According to (4) in Table 125A, the rank order of the three H groups is Con., Dien., and Cui. in level of conceptual ability exhibited in connection with two of the arithmetical tasks, computation and problem solving. That is to say, the Con. subjects were the ablest in accounting mathematically for the computations required in the abstract examples and in explaining satisfactorily solutions of the one-step problems, and the Dien. and the Cui. subjects, in the order named, displayed smaller amounts of these abilities. However, in the third task, the number combinations, the three H groups demonstrated equal degrees of conceptual maturity by using, to about the same extent, thought processes of similar degrees of meaningful abstractness.

The safest inference from these findings will include recognition of the exception just noted and might be worded something like this: In the case of very bright children, the Con. program was the most effective of the three, and the Dien. program was the next most effective, in promoting conceptual maturity in ability to explain the mathematical rationale of computations in abstract examples and solutions of one-step problems.

Some may attach special significance to the fact that the ranking, Con., Dien., Cui., holds in two of the three possibilities; point out that, in betting, the odds, 2 out of 3, are regarded as very favorable indeed, and prefer to disregard the exception and, without qualification, to evaluate the over-all effectiveness of the Con. program as greatest, of the Cui. program as least, and of the Dien. program as intermediate between the other two.

In (5), in which groups of children rated relatively low in brightness, a generalization can be arrived at which is hardly controversial. The Con. L group is given the top position for effectiveness in two comparisons and the second position in the third. The Dien. L group occupies a different place in each comparison, but does have top place in one. The Cui. L group is at the bottom in two comparisons and is in second position in the third. The ranking of the three programs for effectiveness in promoting conceptual maturity in the case of children below average in brightness seems to be: Con., Dien., Cui.

The last conclusion, for (5), as well as both the safer and the less cautious conclusions for (4), are subject to correction if factors other than brightness can be shown to have had material effects on the factual findings.

First, this possibility may be explored for (4). The Con. H group had had 4% more teaching time than the Dien. H group and 12% more than the Cui. H group. (Table 119.) It is doubtful, however, that if the Cui. subjects had had as much instruction as the Con. subjects, they could have improved upon their two very low ratings enough to have challenged the Dien. subjects for second place in the ranking for conceptual maturity, to say nothing of challenging the Con. subjects for first place. In this study such small differences in time have not been found to be so important.

To continue with (4): the Cui. group was estimated to be the least successful of the three H groups in learning arithmetic, but only by 0.15 and 0.14 of a rating compared with the Con. and the Dien. groups. It was also estimated to be the least successful in learning to read, but here the differences between ratings are even smaller (0.3 and 0.9). Nothing here requires that the ranking of the three programs in (4) needs to be altered.

As for (5): the Cui. L subjects had had 3% more instruction than the Dien. L subjects had had, and 4% more than the Con. L subjects. Such disparities are certainly negligible; and, in any case, the Cui. group ranks lowest in all three comparisons for conceptual maturity.

The largest difference between pairs of estimates for achievement amounts to 0.14 of a rating in arithmetic and to 0.13 of a rating in reading. While in each set of estimates the lowest rating was given the Cui. L subjects, it is extremely unlikely that their shortages in these factors could have affected adversely their learning enough to account for much of their shortages in conceptual maturity.

Proficiency

Significant and near-significant differences in proficiency, as determined by work with the interview items only, are summarized for each arithmetical task in Table 125B. As the most trustworthy criterion of proficiency, priority is given to records for Accuracy, and especially for Correct Answers. Measures of Rate have been included in evaluations only when accompanied by high scores for Accuracy.

As was true in the comparisons of H and L groups in Table 125A for level of conceptual maturity, so here the H groups in each of (1), (2), and (3) proved to be more proficient than the paired L groups. All nine comparisons favor the H groups.

Table 125B. Brightness as a Factor. Relative Standing of Groups in Proficiency*

Groups Compared	Number Combinations	Computation	Problem Solving
(1) Cui. H vs Cui. L	H	H	H
(2) Dien. H vs Dien. L	H	H	H
(3) Con. H vs Con. L	H	H	H
(4) Cui. H vs Dien. H vs Con. H	1. Con. 2. Cui. and Dien. equal	1. Con. 2. Cui. and Dien. equal	Equal
(5) Cui. L vs Dien. L vs Con. L	1. Con. 2. Dien. 3. Cui.	1. Con. 2. Cui. 3. Dien.	1. Con. 2. Dien. 3. Cui.

* See the footnote to Table 125A

In the comparisons of the H groups in the three programs, (4), the Con. subjects were the most proficient in two arithmetical tasks, with nothing to choose between the Cui. and the Dien. subjects. In the third task, solving problems, the three H groups were equally proficient.

In the comparisons of the three L groups, (5), the Con. subjects were the most proficient, but this time in all of the arithmetical tasks. The Dien. subjects ranked second in two tasks, and the Cui. subjects ranked second in one.

3. General Arithmetic Achievement as a Factor

It is customary for measures of achievement in the more academic of the elementary school subjects to correlate with measures of brightness to the extent of coefficients of the order, .50 to .65. For this reason, which is supported by the evidence adduced in Chapter 5 for the Scottish study, the findings of comparisons between groups set up, as here, on ratings both for arithmetic and for reading achievement, may be expected to resemble those when groups are constituted on the basis of brightness. Nevertheless, it seems desirable to work through comparisons in terms of one school subject, and achievement in arithmetic has been selected.

It is to be remembered that teachers in rating children for the factor arithmetic achievement had no knowledge of their relative success in the interviews. Instead, they judged their pupils' achievement in arithmetic from what they had observed in their daily performances in the classroom. Moreover, while their observations must have included what their pupils did in connection with arithmetical tasks identical with or similar to those in the interviews, they must have included much more than this. After all, the ratings were for over-all achievement in this subject.

The data in Table 126 show, among other things: (a) about the same total amount of instructional time for the six groups, the largest difference being 14% (in favor of the Con. H group compared with the Cui. H group); (b) average ages within a range of three months; (c) the consistently high ratings, in brightness and in reading achievement, of the groups rated 1 in over-all arithmetic achievement; (d) the consistently much lower ratings, in the same factors, of the groups of subjects rated 3 and 4 in arithmetic achievement; (e) the close similarity of the ratings of the three H groups in brightness and reading achievement; and (f) the larger differences in such ratings for the three L groups.

Table 126. Arithmetic Achievement as a Factor. Composition of the H and the L Groups in Each Program

Bases of Comparison	Cuisinaire		Dienes		Conventional	
	H	L	H	L	H	L
Number of Subjects	72	119	69	184	92	126
Per Cents by Sex						
Male	47	53	49	51	59	52
Female	50	50	52	48	41	48
Weighted Means for						
Total Instructional Time	559	585	569	608	636	584
Age	7 yr. 10 mo.	7 yr. 11 mo.	7 yr. 11 mo.	8 yr. 1 mo.	8 yr. 0 mo.	7 yr. 11 mo.
Estimated Ratings:						
Brightness	1.13	2.90	1.13	3.09	1.16	3.13
Achievement in Reading	1.00	3.28	1.00	3.33	1.00	3.30

A. The influence of wide differences in over-all arithmetic achievement upon the effectiveness of each program.

(1) Cui. H vs Cui. L

Conceptual maturity. (a) The average per cents of the Cui. groups for Rubrics A and B combined are 10 and 51 for the H and L subjects, respectively. The H subjects employed Rational Solutions and Meaningful Habituations, each, more than twice as often as did the L subjects, to secure answers for the eight number combinations. (Table 127.)

Table 127. The Number Combinations. Thought Processes with the Eight Combinations; Averages in Per Cents

Rubrics and Categories	Cuisinaire		Dienes		Conventional	
	H	L	H	L	H	L
A. Non-Attempts	3	16	4	16	1	4
B. Counting	7	35	23	39	15	30
C. Guessing	2	8	3	7	1	10
Rote Recall	6	4	8	4	2	1
Recitation of Tables	13	8	7	7	8	12
Mechanical Uses: Reverses and Inverses	6	6	3	2	1	6
Sub-Totals	27	26	21	20	12	29
D. *Rational Uses:*						
Reverses	1	0	0	0	1	1
Inverses	2	0	4	0	2	3
Solutions, Other Forms*	45	20	38	20	32	19
Sub-Totals	48	20	42	20	35	22
E. Uses of Meaningful Habituations	12	1	7	2	36	13

* The category, Inverses, Plus, is included

The H group has a significantly smaller average per cent than does the L group for Rubrics A + B in work with each pair of combinations. And the H average per cent for Rubric D is significantly the larger for each pair, and for Rubric E,

Table 128. Computation. Quality of Explanations of Computations, Attempts Only; Averages in Per Cents

Classification of Explanations*	Cuisinaire		Dienes		Conventional	
	H	L	H	L	H	L
Satisfactory						
Words Only	24	29	33	13	34	12
Altered Examples	27	17	19	16	42	26
Other	0	3	3	2	0	0
Unsatisfactory						
Words Only	21	17	27	43	13	31
Altered Examples	28	27	7	9	11	28
Other	0	7	10	16	0	2

* *The category Imagery is omitted*

Table 129. Problem Solving. Reasons Given for Solutions, Attempts Only; Averages in Per Cents

Quality of Explanations*	Cuisinaire		Dienes		Conventional	
	H	L	H	L	H	L
Good	67	35	67	46	71	37
Unsatisfactory	10	27	7	24	5	20
Indeterminate	2	5	2	2	4	11

* The category Fair is omitted

for three pairs. The differences in seven of the comparisons amount to 20 percentage points or more.

In ways of thinking answers for the combinations the subjects rated 1 in general arithmetic achievement exhibited a stage of conceptual maturity far in advance of that exhibited by the subjects rated 3 and 4.

(b) In ability to verbalize understanding of the mathematical rationale of the four computations, the H and L groups differ little, if at all. Their average per cents for Satisfactory Explanations are: H, 51; L, 49. (Table 128.)

For the separate computations the per cents for Satisfactory Explanations are:

47 + 46 H, 34; L, 67 3 x 29 H, 61; L, 28
97 - 45 H, 48; L, 9 84 ÷ 4 H, 62; L, 92

In these comparisons each group is seen to have been significantly superior to the other in two examples, a fact which accounts for their closely similar average per cents for Satisfactory Explanations and makes it impossible to give credit to either group for the greater degree of conceptual maturity.

(c) The H subjects greatly exceeded the L subjects in ability to make clear to the interviewers the reasons why they solved the twelve problems as they did. (Table 129.)

The average per cent of the H group for Good Explanations is larger than that of the L group by 27 points or more for each of the four sets of problems.

Proficiency. (a) The H subjects demonstrated by far the greater proficiency in finding answers for the eight combinations. Their margins of superiority are decisive for both categories for Accuracy and for both categories for Rate. (Table 130.)

In like manner they outmatched the L subjects in proficiency with each pair of combinations. For Correct Answers the H group has advantages of at least 12

Table 130. The Number Combinations. Proficiency in Giving Answers: Accuracy and Rate; Averages in Per Cents

Nature of the Measures	Cuisinaire		Dienes		Conventional	
	H	L	H	L	H	L
Accuracy						
Correct Answers	90	60	84	60	94	69
Non-Attempts	3	17	5	17	0	6
Rate, Attempts Only						
Prompt Responses	58	15	50	23	60	28
Long Pauses	7	37	15	34	9	29

Table 131. Computation. Proficiency: Accuracy and Quality of Performance; Averages in Per Cents

Types of Measure	Cuisinaire		Dienes		Conventional	
	H	L	H	L	H	L
Accuracy						
Correct Answers	78	44	69	37	90	62
Non-Attempts	2	21	11	33	0	6
Quality of Performance, Attempts Only						
Smooth	88	65	72	42	91	68
Erratic	11	35	27	58	8	32

Table 132. Problem Solving. Proficiency: Accuracy and Performance Level; Averages in Per Cents

Types of Measure	Cuisinaire		Dienes		Conventional	
	H	L	H	L	H	L
Accuracy						
Correct Answers	93	62	92	73	95	76
Non-Attempts	0	11	2	7	1	3
Level of Performance, Attempts Only						
High	58	27	52	28	78	47
Low	7	46	9	34	7	26

points, and has percentage advantages for Prompt Responses of at least 40 points for all four pairs.

(b) It takes no more than a glance at the H and L average per cents in Table 131 to show that the subjects rated 3 and 4 in general arithmetic achievement were no match for the subjects rated 1 in skill in solving the four examples.

The H per cent for Correct Answers exceeds that of the L group by 28 or more points for each example, and the H per cent for Smooth Performances is the larger of the two by 15 points or more for each example.

(c) The H subjects were also much more proficient in solving the twelve problems than were the L subjects. Each of the four comparisons made possible by the average per cents of the two groups in Table 132 results in a significant advantage for the H group.

For each of the four sets of problems the H group has credit for a percentage advantage of at least 14 points in the category Correct Answers and of at least 25 points in the category High-Level Performances.

(2) Dien. H vs Dien. L

Conceptual maturity. (a) In thought processes used to find answers for the eight combinations the Dien. H group has a significantly smaller average per cent than does the Dien. L group for each of the categories, Non-Attempts and Counting, as well as a significantly larger average per cent for Rational Solutions. Neither group employed Meaningful Habituations to an appreciable extent. (Table 127.)

For each of the four pairs of combinations the L average per cent is substantially the larger of the two for Rubric A and for Rubric B. On the other hand, the H average per cent is substantially the larger of the two for Rubric D for three pairs and for Rubric E for one pair.

The results of the comparisons of average per cents both for the eight combinations as a whole and for the pairs of combinations are in agreement: the H group demonstrated a greater degree of conceptual maturity than did the L group.

(b) The H subjects offered Satisfactory Explanations of the four computations together in 55% of their attempts; the L subjects, in 31%. (Table 128.)

With each of the separate computations the H group exhibited a significantly larger amount of this ability. The difference between the H and the L average per cents for Satisfactory Explanations amounts to 12 or more points in each of the four comparisons.

(c) The H group greatly surpassed the L group in successful attempts to explain solutions of the twelve problems as a whole. (Table 129.)

The average per cents of the two groups for Good Explanations for the sets of problems are as follows:

Addition	H, 86; L, 72	Multiplication	H, 60; L, 40
Subtraction	H, 70; L, 43	Division	H, 54; L, 29

Proficiency. (a) The H subjects, on the average, attempted answers for the eight combinations considerably oftener than did the L subjects, and they were much more successful in furnishing Correct Answers. Moreover, their answers were announced promptly substantially more frequently than were the answers of the L subjects. (Table 130.)

The average per cent of the H subjects for Correct Answers significantly exceeds that of the L subjects for three pairs of combinations (by 28 to 30 points) and almost does so for the fourth (8 points). The H group has the larger average per cent for Prompt Responses for all four pairs, by 25 points or more.

(b) By earning much the larger average per cents for Correct Answers and for Smooth Performances and by having much the smaller average per cent for Non-Attempts, the H subjects showed far more skill than did the L subjects in solving the four abstract examples. (Table 131.)

The greater computational skill of the H subjects was demonstrated in each of the four separate examples, and in much the same manner as above.

(c) The L group was considerably less proficient in problem solving than was the H group. In work with the twelve problems the L group secured many fewer Correct Answers and is credited with many fewer High-Level Performances as well. (Table 132.)

The average per cents of the two groups for Correct Answers for the four sets of problems are:

Addition	H, 97; L, 91	Multiplication	H, 92; L, 73
Subtraction	H, 94; L, 71	Division	H, 86; L, 58

For each set the average per cent of the H subjects for High-Level Performances is the larger of the two by at least 16 points.

(3) Con. H vs Con. L

Conceptual maturity. (a) On the average, the L subjects counted twice as often as did the H subjects to get answers for the eight combinations, used Rational Solutions to a much smaller extent, and employed Meaningful Habituations a little more than a third as commonly, compared with the H subjects. (Table 127.)

For two pairs of combinations the H subjects counted to get answers much less frequently than did the L subjects (there was little counting in either group in the case of the multiplication and division pairs), and for two pairs the H subjects solved rationally and for two other pairs took advantage of Meaningful Habituations much more often than did the L subjects. In a word, whenever the H and L average per cents differ significantly, the difference is in favor of the H group in terms of conceptual maturity of thought processes.

(b) The average per cent of the H group for Satisfactory Explanations of the four computations is double the size of the average per cent for the L group. (Table 128.)

All differences between the H and L average per cents for Satisfactory Explanations of the computations in the four separate examples favor the H subjects by 30 points or more in each instance.

(c) All three comparisons made possible by the data in Table 129 show that the H subjects were far more able than were the L subjects to explain solutions of the twelve problems.

The average per cents of the two groups for Good Explanations of solutions of the four sets of problems are as below.

Addition	H, 79;	L, 54	Multiplication	H, 81;	L, 36
Subtraction	H, 66;	L, 33	Division	H, 60;	L, 25

Proficiency. (a) Compared with the L subjects, the H subjects averaged many more Correct Answers for the eight combinations together, attempted slightly more answers, and about twice as commonly announced answers promptly. (Table 130.)

For all pairs of combinations except that in addition the H group has a significantly larger average per cent for Correct Answers, and for all four pairs without exception, a significantly larger average per cent for Prompt Responses.

(b) The H subjects attempted to solve the four examples in 100% of their opportunities; the L subjects, in 94%. And the H subjects obtained almost a third more Correct Answers than did the L subjects and surpassed the latter by a substantial margin in credits for Smooth Performances. (Table 131.)

In each of the separate examples the H group obtained materially more Correct Answers and has the higher rating in the category, Smooth Performances.

(c) The H average per cents both for Correct Answers and for High-Level Performances are much larger than those of the L group, in attempts to solve the twelve problems. (Table 132.)

The average per cent of the H group for Correct Answers exceeds that of the L group for each set of problems, as below.

Addition	H, 99; L, 92	Multiplication	H, 95; L, 72
Subtraction	H, 93; L, 73	Division	H, 95; L, 66

The record of the H group with respect to High-Level Performances is much better than that of the L group for each set of problems.

B. The relative effectiveness of the three programs when taught to children rated 1 in over-all achievement in arithmetic and when taught to children rated 3 and 4 in this factor.

(4) Cui. H vs Dien. H vs Con. H

Conceptual maturity. (a) According to the summary below, which is based on Table 127, the three groups of subjects attempted answers for the eight combinations to about the same extent, but the Cui. H subjects counted somewhat less often than did the Con. H subjects and much less often than did the Dien. H subjects. The Cui. H group has the largest of the three average per cents for uses of Rational Solutions, by a significant margin over the Con. H group. On the other hand, on the average, the Con. H subjects took advantage of Meaningful Habituations three times as often as did the Cui. H subjects and five times as often as did the Dien. H subjects. The fact that the Con. H group had habituated so many of the combinations (over a third) probably accounts for its relatively smaller average per cent for Rubric D, Rational Solutions.

Rubrics	Average Per Cents		
	Cui. H	Dien. H	Con. H
A. Non-Attempts	3	4	1
B. Counting	7	23	15
C. Guessing, Rote Recall, Recitation of Tables, etc.	27	21	12
D. Rational Solutions	48	42	35
E. Uses of Meaningful Habituations	12	7	36

From these comparisons of averages one might infer that in relative degree of conceptual maturity the three groups clearly rate in the order: Con., Cui., and Dien. This inference has to be changed in part, however, when the records of the groups with pairs of combinations are examined, as below. There is little question that the Dien. group is properly ranked as lowest; but the Cui. group becomes stronger as a contender for first place. Chiefly because of their better showing in Rubric E, the Con. subjects should probably continue to be ranked above the Cui. subjects, although their margin of superiority is considerably reduced.

Pairs of Combinations	Rubrics		
	A + B	D	E
Addition	Cui.** Con.*	Cui.* (Cui.*) Con.*	Cui.* Con.* Dien.
Subtraction	Cui.** (Con.*)	(Cui.*) - -	(Cui.*) - -
Multiplication	(Con.*) - -	Cui.* (Dien.*) Con.	Con.** - -
Division	Cui.* Con.* Dien.	Cui.* Dien.* Con.	Con.** Cui.*

(b) The average per cents of the three H groups for Satisfactory Explanations of the four computations are: Cui., 51; Dien., 55; Con., 76. (Table 128.) Their per cents for this category for the separate computations are, in the order, Cui., Dien., Con.:

27 + 46	34; 65; 81	3 x 28	61; 55; 75
93 - 45	48; 56; 75	84 ÷ 4	62; 48; 72

By significant margins in each instance the Con. subjects exhibited greater ability than did the Cui. and the Dien. subjects to account mathematically for the computation in each example. The Dien. per cent for Satisfactory Explanations exceeds that of the Cui. group significantly for the addition example and nearly significantly for the subtraction example, whereas the Cui. group has a much better record than the Dien. group for the division example and a little better record for the multiplication example.

In explaining computations the Con. subjects displayed much the most advanced stage of meaningful abstraction, and the Cui. and the Dien. groups were a match for each other.

(c) The three groups explained solutions of the twelve problems equally well (Table 129), as they did also solutions of the four sets of problems. Of the twelve comparisons of pairs of average per cents for Good Explanations, only one shows a substantial difference, and none of the other eleven shows a difference as large as 9 points. Quite apart from matters of significance, the Dien. subjects have the largest of the three average per cents for Good Explanations for one set; the Con. subjects, for three sets.

Proficiency. (a) The Con. subjects gave Correct Answers in 94% of their attempts with the eight number combinations as a whole; the Cui. subjects, in 90%; and the Dien. subjects, in 84%. As for Prompt Responses, there is only one significant difference between the average per cents of the two groups in each pair, and it is in favor of the Con. group over the Dien. group. (Table 130.) In both of these categories the Con. group surpassed the Cui. group, but the margins are only 5 and 2 percentage points, respectively.

The Con. subjects were more accurate than the Dien. subjects for three pairs of combinations, by 9 percentage points for two pairs and by 14 points for the third. Likewise, the Con. subjects were more accurate than the Dien. subjects for three pairs, by only 5 percentage points for two pairs and by 14 points for the third.

In proficiency with the number combinations the Con. subjects rate highest, the Cui. subjects somewhat lower, and the Dien. subjects considerably the lowest.

(b) In skill in solving the four examples combined the Con. group has much the largest average per cent for Correct Answers, and the Cui. group the next highest, with an advantage of 9 points over the Dien. group. The last-named group made no attempt to solve the examples in 11% of their opportunities. The Con. and the Cui. groups have about the same record for Smooth Performances, and both of these groups considerably surpass the Dien. group in this category. (Table 131.)

For each of the four examples the Con. group has the largest per cent for Correct Answers. In seven of the eight comparisons with the other two groups, the advantage for the Con. group is substantial. In only one of the four comparisons with the Dien. group, which regularly was the least accurate of the three groups, the Cui. group has a significant advantage. The Con. group also has consistently the best of the three records for Smooth Performances; the Dien. group, the poorest.

(c) The three groups showed about the same degree of skill in solving the twelve problems, at least as far as Accuracy is concerned, though the Con. group is credited with considerably more High-Level Performances than is either of the other groups.

The average per cents of the three groups for Correct Answers for the sets of problems, in the usual order, are:

Addition	93; 86; 95	Multiplication	90, 92; 95
Subtraction	92; 94; 93	Division	93; 86; 95

There is nothing in the set-by-set comparisons to require a change in the evaluation of the three groups as being equal in ability to solve one-step problems.

(5) Cui. L vs Dien. L vs Con. L

Conceptual maturity. (a) The summary below shows, for the eight number combinations, the average per cents of the three L groups for the five rubrics of thought processes. The data are from Table 127.

Rubrics	Average Per Cents		
	Cui. L	Dien. L	Con. L
A. Non-Attempts	16	16	4
B. Counting	35	39	30
C. Guessing, Rote Recall, Recitation of Tables, etc.	26	20	29
D. Rational Solutions	20	20	22
E. Uses of Meaningful Habituations	1	2	13

The significant and near-significant differences in comparisons of average per cents for the pairs of combinations are as follows.

Pairs of Combinations	Rubrics		
	A + B	D	E
Addition	Cui.**	Con.* (Con.*)	- -
	Con.*	- -	- -
Subtraction	Con.**	Con.**	- -
	Cui.*	- -	- -
Multiplication	Con.**	Dien.**	Con.**
	(Cui.*)	Cui.*	- -
Division	Con.**	- -	Con.**
	(Cui.*)	- -	- -

While the greater use made of Meaningful Habituations and of Rational Solutions of the Con. L subjects is restricted in each case to two pairs, their record for Rubrics A + B is relatively so good that they should be rated highest in conceptual maturity, and the Cui. L and the Dien. L subjects have about equal claims to be rated next highest, for, if the Cui. L subjects seem to have surpassed the Dien. L subjects in conceptual maturity, they did so by too narrow a margin to warrant giving them undisputed possession of second place.

(b) The Cui. group was substantially the most successful of the three in explaining the four computations as a whole, and the Con. group was somewhat more successful than was the Dien. group. (Table 128.)

For the separate computations the per cents for Satisfactory Explanations are as below. Those for the Cui. group come first; those for the Con. group, last.

27 + 46	67; 43; 42	3 x 28	28; 28; 33
93 - 45	9; 37; 37	84 ÷ 4	92; 19; 42

The record of the Cui. group is especially interesting, for with by far the largest per cents for the addition and multiplication examples, it has by far the smallest for the subtraction example. The records of the Dien. and the Con. groups differ little except in the case of the division examples, for which the latter group has much the larger per cent.

All in all, the Cui. subjects clearly demonstrated the most ability to verbalize their understanding of the mathematics of computations, and so win first place in the ranking for conceptual maturity. The Con. subjects have second place, but by no considerable margin over the Dien. subjects.

(c) The Dien group has an 11-point advantage over the Cui. group in average per cents for Good Explanations of the solutions of the twelve problems, and has a 9-point advantage over the Con. group. (Table 129.)

For each set of problems the Dien. subjects have the largest average per cent for Good Explanations. Of the twelve comparisons of such per cents, the

Dien. subjects have significant advantages in four (all in the addition and subtraction sets) and a near-significant advantage in a fifth. In only one of the remaining seven comparisons is the difference as large as 6 points.

As indicated by relative ability to explain solutions of problems, the Dien. group displayed a level of conceptual maturity higher than that of either of the other two groups. The Con. and the Cui. groups were about a match for each other, with a slight edge in favor of the Con. group.

Proficiency. (a) For the eight combinations, the average per cents of both the Cui. and the Dien. groups are 11 points larger than that of the Con. group for Non-Attempts, and both average per cents are 9 points smaller for Correct Answers. The Con. subjects are credited with considerably more Prompt Responses than are the Cui. subjects. (Table 130.)

Significant and near-significant differences between average per cents for Correct Answers for the pairs of combinations are:

Addition: Dien., 1 near-significant

Subtraction: Dien., 1 near-significant

Multiplication: Con., 2 significant

Division: Con., 1 significant; 1 near-significant.

Of the twelve chances for the average per cent of one group to exceed that of another for Prompt Responses, six are in favor of the Con. group, and one is in favor of the Dien. group.

On the record as a whole, the Con. subjects showed the most proficiency with the number combinations, followed by the Dien. subjects, then by the Cui. subjects.

(b) The Con. group was far more accurate than was each of the other two groups in computing answers for the four examples. The average per cent of the

Con. subjects is much the largest of the three for Non-Attempts. In the category, Smooth Performances, the Con. subjects greatly surpassed the Dien. subjects, who consistently made the poorest record for proficiency. (Table 131.)

The Con. subjects have much the largest per cent for Correct Answers for each of the examples, and the Cui. subjects have much larger per cents for this category than do the Dien. subjects for two examples. As for Smooth Performances, the Dien. group consistently has the smallest per cent of the three groups, and the Cui. and the Con. groups each have the larger per cent for two examples.

(c) The Dien. and the Con. groups have very similar records for Accuracy (Correct Answers and Non-Attempts) in solving the twelve problems, and their records are considerably better than the Cui. record for the category, Correct Answers. The Con. group clearly surpassed the other two groups in frequency of High-Level Performances. (Table 132.)

The average per cents of the three groups for Correct Answers for the sets of problems are as follows, in the order Cui., Dien., Con.

Addition	83; 91; 92	Multiplication	46; 73; 72
Subtraction	67; 71; 73	Division	53; 58; 61

The Con. subjects have the largest average per cent for High-Level Performances for all four sets, and it is significantly the larger in seven of the eight comparisons with the Cui. and the Dien. subjects.

Summary, Section 3

Conceptual maturity

Table 132A summarizes the significant and near-significant advantages of the different groups with respect to level of conceptual maturity displayed in dealing with each of the arithmetical tasks.

Table 132A. Arithmetic Achievement as a Factor.
Relative Standing of Groups in Conceptual Maturity*

Groups Compared	Thought Processes with Combinations	Ability to Explain Computations	Ability to Explain Solutions of Problems
(1) Cui. H vs Cui. L	H	Equal	H
(2) Dien. H vs Dien. L	H	H	H
(3) Con. H vs Con. L	H	H	H
(4) Cui. H vs Dien. H vs Con. H	1. Con. 2. Cui. 3. Dien.	1. Con. 2. Cui. and Dien. equal	Equal
(5) Cui. L vs Dien. L vs Con. L	1. Con. 2. Cui. and Dien. equal	1. Con. 2. Cui. 3. Dien.	1. Dien. 2. Cui. and Con. equal

* In (4) and (5) no distinction has been made between differences that are significant and those that are slightly less than significant.

Intra-program comparisons. The groups rated 1 in over-all achievement in arithmetic made a better showing than did their paired groups of children rated 3 and 4 in this factor, with respect to conceptual maturity in work with all three arithmetical tasks in comparisons (2) and (3); and in (1) the H group made a better showing than the L group in work with two of these tasks.

In no case could the results have been affected by differences in amounts of instructional time, for in (3) only had the H group had more instruction than the L group, and then by only 9%.

On the other hand, the general superiority of the H groups in conceptual maturity cannot be ascribed purely to their being better achievers in arithmetic alone, as judged by teachers, for in each of (1), (2), and (3) the H subjects were much brighter than their paired L subjects and much more successful in reading. On the average, the children in the three H groups were estimated to be brighter by 1.90 and better readers by 2.30 than the L subjects; and these differences on the 4-point rating scale used are very large indeed.

On this account, the greater conceptual maturity of the H groups should almost certainly be attributed, not to the greater ability the H subjects had demonstrated in learning arithmetic in general, but to the simple fact that they were much superior to the L subjects in learning ability.

Cross-comparisons of programs. In comparison (4) the three groups involved were made up of the children who, in each program, were rated 1 in over-all achievement in arithmetic.

According to the entries in Table 130A, the Con. H group, on the average, used ways of thinking answers for the number combinations that represented higher levels of meaningful abstractness than did those of the Cui. and the Dien. H groups and were more successful than they in accounting mathematically for computations in abstract examples. They were able to do no better, however, than the subjects in the other two groups in giving Good Explanations of solutions of one-step problems. The Cui. H subjects employed thought processes with the combinations that were slightly more mature than were those of the Dien. H subjects, but just equalled the latter subjects in ability to explain computations and solutions of problems.

On the basis of these factual findings, the Con. program was the most effective of the three in promoting conceptual maturity in the case of excellent achievers in arithmetic, and the Cui. program was next most effective, though but little more so than was the Dien. program.

But is this ranking based solely upon the comparative merits of the three programs as such, or is it the result of differences among the groups with regard to other factors? True, the Con. H subjects had the benefit of more teaching time than did the subjects in the other groups, -- 12% more than those in the Dien. H group and 14% more than those in the Cui. H group. On the other hand, these disparities in amount of instruction are probably too small to have given the Con. H group much of an advantage. The teacher ratings of the three groups are identical in reading achievement and nearly identical in brightness. (Table 126.) If the Con. H subjects were favored in other ways, they are unknown. In the absence of information relevant to this matter, one must apparently assume that, as the three programs were taught in the English schools, the Con. program was actually the most effective of the three according to the criteria used in this study.

In comparison (5), the children in the groups representative of the three programs were rated 3 and 4 in achievement in arithmetic in general.

To refer again to Table 132A: the children in the Con. L group displayed higher levels of meaningful arithmetical abstractions than did the children in the Cui. and the Dien. L groups, in the manner in which they dealt with the number combinations and the computations. However, they were able just to equal the two other groups in this respect, in explaining solutions of problems. All things considered, the Con. program was the most effective of the three, with nothing to choose between the Cui. and the Dien. programs for second place in the ranking. This evaluation must of course be regarded as tentative until it is determined whether factors other than those within the programs themselves influenced their relative effectiveness.

If there were such factors, amount of instruction was not one of them, for the Con. L group had had the same total of teaching time as the Cui. L group and had had 4% less than the Dien. L group. The Cui. L subjects have a rating in brightness 0.19 higher than that of the Dien. L subjects and 0.23 higher than that of the Con. L subjects, and the ratings of the three L groups in reading achievement differ scarcely at all. (Table 126.)

So, as in comparison (4), one is left in the position of having to assume that, save for the possible effects of unknown factors (if such there be), the Con. program was more successful than the Cui. program and the Dien. program, in advancing the conceptual maturity of poor achievers in arithmetic, and that these two last-named programs were equally effective in this regard.

Proficiency

Table 132B is a summary of the results of comparisons of groups with reference to proficiency in working with the interview items in the three arithmetical tasks. Only significant and near-significant differences are included.

Table 132B. Arithmetic Achievement as a Factor.
Relative Standing of Groups in Proficiency*

Groups Compared	Number Combinations	Computation	Problem Solving
(1) Cui. H vs Cui. L	H	H	H
(2) Dien. H vs Dien. L	H	H	H
(3) Con. H vs Con. L	H	H	H
(4) Cui. H vs Dien. H vs Con. H	1. Con. 2. Cui. 3. Dien.	1. Con. 2. Cui. 3. Dien.	Equal
(5) Cui. L vs Dien. L vs Con. L	1. Con. 2. Dien. 3. Cui.	1. Con. 2. Cui. 3. Dien.	1. Dien. and Con. equal 3. Cui.

* See footnote to Table 132A

In comparisons (1) to (3), the H group in each program was found always to be more proficient than its paired L group, just as was the case (with one exception) in the corresponding comparisons for conceptual maturity. And the explanation given for the superiority of the H groups in conceptual maturity seems to serve equally well for their superiority in proficiency. While the H subjects, by design, greatly excelled in over-all achievement in arithmetic (the basis on which the H and L groups were constituted), they also greatly excelled in brightness and in demonstrated ability to learn to read. The marked differences in favor of the H subjects in these factors -- probably the reason why they had earned their higher rating for arithmetic achievement, to begin with -- were, in all likelihood, the reason why they were more proficient than the L subjects in this inquiry.

In comparison (4), for groups of subjects rated 1 in arithmetic achievment (Table 132B), the Con. H group was the most proficient of the three H groups in two of the three arithmetical tasks, and the Cui. group next most proficient.

In comparison (5), for groups of subjects rated 3 and 4 in arithmetic achievement, the Con. group is ranked first in proficiency, and the Cui. and the Dien. groups are about equally matched for second place.

Scrutiny of the data collected in this investigation failed to disclose any explanation for the top ranking of the Con. H and the Con. L groups in comparisons (4) and (5), respectively; that is to say, any explanation other than the Con. program itself. Until some other explanation is found, one is left with the conclusion that, of the three programs as they were taught in English schools, the Con. program was the most successful in developing proficiency, in the limited sense in which that term is used here.

Chapter Summary

Questions A and B

The purpose of this chapter was to examine the possible influence of three of the four factors* considered in the Scottish study on the effectiveness and the relative effectiveness of the three programs in promoting conceptual maturity. The findings of comparisons in Sections 1 to 3 of the chapter will be brought together and organized around two questions, A and B.

A. Comparisons (1) to (3) were planned to determine how, if at all, the effectiveness of each program varied when it was taught to children rated 1 and when taught to children rated 3 and 4 in each factor.

Time. In comparisons (1) to (3), the H groups, as hoped for, had had at least 52% more teaching time than the L groups had had. Nevertheless, in only two of the nine comparisons (three arithmetical tasks, three paired groups) did the H group exhibit substantially the larger degree of conceptual maturity, while in three (in two by significant margins) superiority lay with the L groups. The inference would seem to be that marked differences in amount of instruction had little, if any, effect on conceptual learning.

Yet, in (1), for the Cui. program, the H subjects were handicapped by being much less bright than the L subjects and much less successful in learning arithmetic in general and in learning to read, disadvantages which clearly could have offset their advantage in amount of instruction. Hence, the results in comparison (1) are irrelevant to the question at issue.

---

* The omitted factor is Reading Achievement, for which the basic tables only will be found in the Appendix. For the reason for this omission, see the first page of this chapter.

By contrast, the results of comparisons (2), for the Dien. groups, and (3), for the Con. groups, are distinctly relevant, for in both, in addition to their greater amounts of teaching time, the H subjects rated higher in brightness than the L subjects and were regarded as better achievers in arithmetic and reading, in one or both; and yet they failed to show superiority in conceptual maturity. The findings in (2) and (3) therefore support the initial inference to the effect that, by themselves, differences in teaching time were not, by themselves, decisive in determining advancement in conceptual learning.

Brightness. Marked disparities in brightness on the part of the children taught were exceedingly influential in determining the effectiveness of each of the programs. The evidence on this point is overwhelming, for in all nine of the comparisons in (1) to (3) the H groups revealed unmistakably higher levels of meaningful arithmetical abstractions than did the L groups.

As has been said before, "it is hardly astonishing" that in each program the children rated 1 should greatly surpass in conceptual maturity those rated 3 and 4 in brightness; and, other things being equal, anything else would be difficult to explain. Of course the H subjects regularly rated much higher than the L subjects in achievement in arithmetic and in reading; but these advantages are typically accompaniments, if not consequences, of superior brightness, and they should not be allowed to lessen the importance attached to brightness as such as a factor in influencing the effectiveness of the three programs.

Over-all achievement in arithmetic. In eight of the nine comparisons, the subjects rated 1 in over-all achievement in arithmetic revealed significantly higher degrees of conceptual maturity than did the subjects rated 3 and 4 in this factor. These findings would seem to establish beyond question that relative arithmetic achievement was decisive in influencing the effectiveness of each of the three programs. Stated differently, they would seem to mean that the Cui. program, for example, was more successful in promoting conceptual maturity when taught to excellent achievers in arithmetic in general than when taught to poor achievers, and just because of the differences in this factor alone.

Such a conclusion is untenable, however. While the H and the L groups in each program had had about the same amounts of teaching time, they differed greatly in other important respects. The average rating of the three H groups are 1.4 in brightness and 1.00 in reading achievement. The average ratings of the three L groups are 3.04 in brightness and 3.30 in reading achievement. The implication of these comparisons is that the H subjects were superior to the L subjects in conceptual maturity, not because they were better achievers in arithmetic, but because they were much brighter to start with and were more able to profit from school instruction.

B. Comparisons (4) and (5) should produce results helpful in determining the relative effectiveness of the three programs when all were used with children rated 1 and when all were used with children rated 3 and 4 in each of the three factors.

Time. Comparison (4) involves Cui., Dien., and Con. groups which had had relatively large amounts of teaching time. The Dien. H group had had only 4% more instruction than the Con. group and only 6% more than the Cui. group. These differences are too small to have affected the results to any appreciable extent.

According to Table 118A, the Con. group displayed significantly higher degrees of conceptual maturity than did the Cui. and the Dien. groups in work with all three arithmetical tasks, and the Cui. and the Dien. groups exhibited about the same degree of conceptual maturity.

However, the ratings of the Con. H group in brightness and in school achievement are much larger than those of the other two H groups. On this account it is questionable whether, under conditions of approximately equally large amounts of instruction, the Con. program was actually as much more effective than the other programs as first seemed to be the case.

In comparison (5), for the L groups in the three programs, the Con. program, according to the entries in Table 118A, seems to rate first in effectiveness in promoting conceptual maturity, the Dien. program second, and the Cui. program third.

There is no reason to question the position of the Cui. program in this ranking; but not so for the relative rankings of the Con. and the Dien. programs. Besides having had 14% more instruction than the Dien. subjects, the Con. subjects were estimated to be brighter (by 0.24 of a rating) and to be better readers (by 0.16 of a rating). It is probable that the combination of these advantages for the Con. subjects means that the Con. and the Dien. programs should be rated as equal in effectiveness, -- this, under conditions of relatively small amounts of teaching time.

Brightness. Comparison (4) involves Cui., Dien., and Con. groups rated 1 in brightness. As is shown in Table 125A, the three H groups displayed about the same level of conceptual maturity in ways of thinking answers for the number combinations, but in work with the other two arithmetical tasks the Con. H group exhibited higher levels of meaningful arithmetical abstractions than did the other two H groups, and the Dien.H group, than did the Cui. H group. No differences of consequence in other factors, such as in amount of instruction, could have affected the results appreciably, so that if one accepts

2 out of 3 chances as good odds, the three programs rate Con., Dien., and Cui., for effectiveness in promoting conceptual maturity in the case of very bright children.

In comparison (5), the children in the three groups were rated 3 or 4 in brightness, and the Cui. L, the Dien. L, and the Con. L groups differed little with respect to relative amounts of instruction and to ratings in school achievement.

While none of the L groups surpassed the other two L groups in the degree of conceptual maturity exhibited in work with all three arithmetical tasks, the Con. L group showed superiority over both the other L groups in two of the tasks, and the Dien. L group, in one. (Table 125A.) It is quite safe to conclude that, with children rated as below average in brightness, the Con. program was the most effective in promoting conceptual maturity, and the Dien. program next most effective.

Arithmetic achievement. The findings of comparisons (4) and (5), for H groups and for L groups, respectively, have been so recently summarized and interpreted that they will be but briefly mentioned here.

According to Table 132A, the ranking of the programs for effectiveness in teaching children rated 1 in over-all achievement in arithmetic is: Con., Cui., Dien., though the difference between the two named last is small. And in the ranking of the programs in the case of children below average in arithmetic achievement the Con. program takes first place, and the Cui. and the Dien. programs are about equally qualified candidates for second place. Nothing could be found by way of differences in amounts of instruction or of differences in ratings to require any changes in the rankings as listed above.

Additional data on the relative effectiveness of the programs

It is possible to derive from the entries for comparisons (4) and (5) in Tables 118A, 125A, and 132A a single consolidated "score," as it were, for each of the programs for effectiveness in promoting conceptual maturity. The method is crude, however, and, among other things, requires that the influence of factors other than the one used to set up the groups in each particular table be disregarded. For example, in Table 118A in Section 1 amount of instruction was the factor in question, and in the accompanying text the effects of differences in the factors, brightness and school achievement, were considered in interpreting the results of comparisons (4) and (5). This will not be done here.

The first step is to summarize the frequencies with which the groups in each program earned the ranks 1, 2, and 3. In each of the six (4) and (5) comparisons, ranks were assigned in the tables mentioned above for the degree of conceptual maturity shown in each of the three arithmetical tasks. In the tabulation which follows the first row of data is read thus: According to Table 118A, in which the factor was Time,

> the rank of 1 was won three times by the Con. H group (and so, was won not at all by the Cui. H and the Dien. H groups),
>
> the rank of 2 was twice shared by the Cui. H and the Dien. H groups and was once won by the Dien. H group,
>
> the rank of 3 was assigned only once (because of the two ties for rank 2), and it went to the Cui. H group.

Table, Factor, Comparison	Ranks, with Frequencies		
	1	2	3
Table 118A: Time			
(4) H Groups	Con. (3)	(Cui. Dien.) (2) Dien. (1)	Cui. (1)
(5) L Groups	Dien. (1) Con. (2)	Cui. (1) Dien. (1) (Cui. Con.) (1)	Cui. (1) Dien. (1)
Table 125A: Brightness			
(4) H Groups	(Cui. Dien. Con.) (1) Con. (2)	Dien. (2)	Cui. (2)
(5) L Groups	Dien. (1) Con. (2)	Cui. (1) Dien. (1) Con. (1)	Cui. (2) Dien. (1)
Table 132A: Arithmetic Achievement			
(4) H Groups	(Cui. Dien. Con.) (1) Con. (2)	Cui. (1) (Cui. Dien.) (1)	Dien. (1)
(5) L Groups	Dien. (1) Con. (2)	Cui. (1) (Cui. Dien.) (1) (Cui. Con.) (1)	Dien. (1)

It is of interest to stop long enough to total the number of times the groups representing the three programs have undisputed claims to the various ranks:

Rank 1: Cui., 0; Dien., 3; Con., 9

Rank 2: Cui., 4; Dien., 5; Con., 1

Rank 3: Cui., 6; Dien., 4; Con., 0

In this tabulation the Con. groups have by far the largest number of rankings for first place and never ranked lower than second place, and then but once. By contrast, the Cui. groups never earned first place and ranked in third place six times.

To return to the problem of calculating "scores" for the three programs with reference to their relative effectiveness in promoting conceptual maturity, - suppose that values be assigned to ranks in the following manner:

Rank 1, 4 points   Rank 2, 2 points   Rank 3, 0 points

In the case of ties for a given rank, the values can be divided among the groups involved like this:

A two-way tie for rank 1, and therefore for rank 2 also:

(4 + 2) ÷ 2; so, 3 points to each group;

A three-way tie for rank 1, and therefore for ranks 2 and 3 as well:

(4 + 2 + 0) ÷ 3; so, 2 points to each group;

A two-way tie for rank 2, and therefore for rank 3 also:

(2 + 0) ÷ 2; so, 1 point to each group.

Application of this scheme resulted in "scores" for successfully encouraging children to move steadily to higher levels of meaningful arithmetical abstractions as follows:

Cui., 18; Dien., 30; Con., 52.

To the extent that the rough procedures employed are acceptable, in the investigation in English schools the Con. program was far and away the most effective of the three, and the Cui. program was the least effective, with the Dien. program in a position intermediate between the other two.

Conceptual maturity vs proficiency

All the data in Tables 118A, 118B, 125A, 125B, 132A, and 132B are available for the study, for the problem is not the comparative effectiveness of the programs, but the relationship between degree of conceptual maturity on the one hand and degree of proficiency in work with the interview items on the other.

The nine intra-group comparisons of the types (1), (2), and (3) afforded a total of 27 chances for one group, H or L, to earn a higher rating than its paired group for conceptual maturity. In 22 instances, one group did surpass the other by significant or near-significant margins. (Tables 118A, 125A, and 132A.) The group with the higher rating in conceptual maturity excelled in proficiency in nineteen instances, equalled that group in proficiency in one, and was inferior in two.

The cross-program comparisons of the types (4) and (5) showed fifteen cases in which one of the groups ranked highest in conceptual maturity. Eleven times this group also ranked 1 in proficiency, and three times it ranked 2 or was tied with another group for this position. On the other hand, not once did the group lowest in conceptual maturity rate highest in proficiency. Three times it was the least proficient of the three groups, and six times it was tied for second place.

The evidence as a whole is convincingly clear that progress in conceptual maturity, instead of slowing down progress in proficiency, may well have had the opposite effect.

PART IV
FINAL CHAPTER

CHAPTER 8

FINDINGS, INTERPRETATIONS, IMPLICATIONS, AND RELATED RESEARCH

Findings

This investigation was undertaken as an attempt to evaluate certain instructional programs in arithmetic with respect to their relative effectiveness, over a period of three years, in promoting conceptual maturity on the part of children subjected to them. This purpose seems simple enough, for the scope of the study was very restricted. Comparisons were to be made on a very narrow base, namely, the relative levels of meaningful arithmetical abstractions which had been developed in the case of three common types of arithmetical tasks.

If the purpose of the investigation had been to evaluate the programs in question as wholes, -- that is to say, to determine their relative success in teaching the total content of arithmetic assigned to the first three years, and in all dimensions of the learning, -- in these circumstances one might have predicted trouble. For the research required would have had to be of formidable proportions.

But this was not the purpose of the present inquiry. As the problem has been defined above, it seems to be quite uncomplicated; and decisive findings might reasonably be expected. One might anticipate the data to show unmistakably that Program A was more effective than was Program B (and than was Program C). No such definitive conclusion has eventuated, however. Instead, the findings on the major problem are in conflict, a fact which could imply that the problem was not, after all, so simple as might have been supposed.

In the following short summary, the findings are reported factually, without comment, and many of the supporting data must perforce be omitted. For details, the reader is referred to Parts II and III.

The relative effectiveness of the programs

In the Scottish schools, where the results of instruction by the Cuisinaire (Cui.) program and by the Conventional (Con.) program are compared, the Cui. program was found to be much the more effective of the two in promoting conceptual maturity. The superiority of the Cui. program is made the more impressive by reason of the fact that the Con. subjects, on the average, had had about 23% more instructional time in arithmetic than the Cui. subjects had had (Table 1), and they were rated as somewhat brighter than the Cui. subjects and as better achievers in arithmetic in general and in reading (Table 3).

Their advantages in amount of instruction or in brightness or in demonstrated ability to learn two of the more academic school subjects could conceivably have enabled the Con. subjects to demonstrate the higher status with reference to meaningful arithmetical abstractions; and the combination of their advantages in all these factors might well have produced this effect. That the Cui. subjects overcame their handicaps and displayed the higher status is viewed as evidence that, as the two programs were taught in the Scottish schools, the Cui. program must have had certain features favorable to the promotion of conceptual maturity that were absent or very much less prominent in the Con. program.

In the English schools, comparisons were made of the results obtained by the teaching of three programs, the Dienes (Dien.) program in addition to the Cui. and the Con. programs; and the order of their effectiveness in promoting conceptual maturity proved to be: Con., Dien., Cui. In a word, the relative positions of the Cui. and the Con. programs in the Scottish schools were reversed in the English schools; and the "new" Dien. program, taught only in English schools at the time, was more effective than the Cui. program, but not so consistently and so decisively as is true in the comparison Cui. vs Con.

The evidence supporting the ranking of the three programs in the English schools is most completely assembled in the data for cross-program comparisons (4) and (5) in Table 118A, 125A, and 132A; and the margins of the differences among the programs are shown by their "scores" as derived from the sources just mentioned: Con., 52; Dien., 30; Cui., 18.

No explanation for the reversal of the positions of the Cui. and the Con. programs or for the intermediate position of the Dien. program is to be found in grossly dissimilar amounts of teaching time. As a matter of fact, the Con. subjects had had only 9% more instruction, on the average, than the Cui. subjects had had, and only 7% more than the Dien. subjects had had. (Table 101.) These differences in favor of the Con. subjects are negligible and could have affected the results little, if at all.

According to teacher estimates (Table 103), the Con. subjects enjoyed advantages over the Cui. and the Dien. subjects as follows:

> in brightness, by 0.06 and by 0.18 of a rating;
>
> in over-all arithmetic achievement, by 0.10 and by 0.26 of a rating;
>
> in reading achievement, by 0.13 and by 0.17 of a rating.

The differences in these factors for the Con. and the Cui. groups, none larger in amount than 0.13 of a rating, can probably be safely dismissed as accounting for more than a small fraction of the large superiority of the Con. subjects in conceptual maturity.

In the ratings above, the Dien. subjects apparently were the least blessed of the three groups in brightness and in success in school learning. In comparisons with the Cui. subjects, they were rated lower in each of the three factors, by margins of from 0.04 to 0.16 of a rating. While the disparities in their favor are all small, at least it can be said that the Cui. subjects had no handicaps in learning ability and on this account could have been expected, other things being equal, to show the same degree of conceptual maturity as did the Dien. subjects. Yet, they did not.

Again, one is left with a qualified conclusion: As the three programs were taught in the English schools, the Con. program was the most effective, and the Cui. program the least effective, in promoting conceptual maturity.

Factors possibly influencing the results

There being the chance that the ranking of programs with regard to effectiveness was attributable, not to features of the programs themselves, but to the influence of factors extraneous to them, a study was made to ascertain the reality of this possibility. Five factors which could have affected the results were identified: amount of instruction, brightness, overall achievement in arithmetic, reading achievement, and age. The last-named factor was not open to investigation since groups of sufficient size that differed materially in average age could not be set up for comparisons. The first four factors were examined in the Scottish study; only the first three, in the English study.

Amount of instruction. In the Scottish study, differences in totals of teaching time rarely could be proved to have had any effect on the results of comparisons of the type Cui. H vs Cui. L* and of the type Cui. H vs Con. L.

---

* The letter H designates groups which had a supposed advantage in the factor in question (here, amount of instruction); the letter L, groups which were disadvantaged with respect to the same factor.

In some instances the amounts of instruction of the H and L groups in a pair were about the same; in others, when the subjects in the H group exhibited the higher degree of conceptual maturity, the probable cause was their greater learning ability; and in still others, the supposed L group had already had more instruction than the paired H group.

The Cui. H and the Con. H groups, as desired, had had about the same large amount of instruction, and under these conditions the Cui. program was found to be more effective than the Con. program. But it is not known which, if either, of the programs was more effective under conditions of equally small amounts of instruction, because this condition was not met. The Con. L subjects had had 40% more teaching time than the Cui. L subjects had had.

The H group in each of the programs (Cui., Dien., and Con.) in the Scottish schools regularly had had substantially more instruction than the L group in the same program had had. Yet, with but a few questionable exceptions, these differences in teaching time had little or no effect on the degree of conceptual maturity that had been achieved by the H subjects and by the L subjects when exposed to the same program.

Under conditions of equally large amounts of instruction the Con. program was the most effective of the three, and the Cui. and the Dien. programs were evenly matched in this respect. Under conditions of equally small amounts of instruction, the Con. and the Dien. programs were tied for first place in effectiveness, with the Cui. program far behind.

Brightness. The subjects in the H groups were rated 1 in brightness; those in the L groups, 3 and 4. Hence, there were marked differences in this factor between the H and the L group in each pair compared, and all the differences were of course in the desired direction, as was not true in the

study of the factor, amount of instruction.

In all comparisons of the types Cui. H vs Cui. L and Cui. H vs Con. L in the Scottish study, the H groups displayed the higher degree of conceptual maturity, thus establishing the fact, not unexpectedly, that brightness was a major factor in determining the effectiveness of the Cui. and of the Con. programs.

The Cui. program was more effective than the Con. program both when the children taught rated high in brightness and when they rated low in this factor.

In the English study each of the three programs was more effective when it was used with children rated 1 than when it was used with children rated 3 and 4.

For effectiveness in teaching both the brighter and the less bright children, the ranking of the three programs is: Con., Dien., and Cui. There is little question that the Cui. program deserved the lowest rating; but in both rankings the Con. subjects surpassed the Dien. subjects in the conceptual maturity they manifested in but two of the three arithmetical tasks.

It should be added that, in the Scottish and also in the English schools, the H groups were estimated to be very much better achievers than the L groups both in arithmetic and in reading. The H groups may have profited from these scholastic advantages, but they (the advantages) are probably best viewed as the usual concomitants, if not the results, of their greater brightness.

Over-all achievement in arithmetic. In terms of the results of factual comparisons only, the Cui. H and the Con. H groups in the Scottish schools were conceptually more mature than their paired L groups in the same program, in dealing with all three arithmetical tasks; and the H groups in the comparisons, Cui. H vs Con. L and Cui. L vs Con. H, were conceptually more mature than were the L groups in working with at least two of the tasks. Thus,

arithmetic achievement would seem to have been an important factor in determining relative degrees of conceptual maturity. This inference, however, is almost certainly unsound because, with great regularity, children rated 1 in arithmetic achievement were also estimated to have much more learning ability than those rated 3 and 4, and this fact furnishes the preferable explanation of their superior showing.

Despite handicaps in relative amounts of instruction, the H subjects achieved a higher degree of conceptual maturity when taught according to the Cui. program rather than according to the Con. program -- this, in connection with all three arithmetical tasks. Likewise, children rated 3 and 4 in arithmetic achievement fared better in this respect in two of the tasks when taught according to the Cui. program.

The findings in the English study in comparisons of H and L groups in the same program are similar to those in the Scottish study, with the H groups in eight out of nine chances surpassing the L groups in conceptual maturity. In the English study, too, the superior showing of the H subjects was caused, not by their being much better achievers in arithmetic, but by their possessing considerably more learning ability.

In teaching children rated 1 in arithmetic achievement, as well as in teaching children rated 3 and 4 in this factor, the rank order of the three programs for effectiveness is: Con., Cui., Dien., with, however, little real difference between the two named last.

Reading achievement. Reading achievement was studied as a factor only in the Scottish schools. (The basic tables for this factor in the English schools are supplied in the Appendix.)

Reading achievement turned out to be a factor only in the invalid sense that made arithmetic achievement a factor, namely, the better showing of the

H groups for conceptual maturity in the results of purely factual comparisons. The teacher ratings of the H and L groups, classified according to reading achievement and taken from Table 33, are:

		Reading Achievement	Brightness	Arithmetic Achievement
Cuisinaire --	H	1.00	1.25	1.00
	L	3.00 or 4.00	3.38	3.48
Conventional -	H	1.00	1.34	1.00
	L	3.00 or 4.00	3.17	3.39

The consistently greater conceptual maturity displayed by the H subjects is less well accounted for by their being much better readers than by their being much better learners. For this reason, the more detailed findings are not presented here.

Conceptual maturity vs proficiency

For the findings with reference to the relationship between level of conceptual maturity and level of proficiency (on the interview items only), the data are taken from Chapter 5 for the Scottish study and from Chapter 7 for the English study. These tables, being more detailed and reporting more analyses than do the corresponding Chapters 4 and 6, contain the results of the larger numbers of relevant comparisons.

In the Scottish study there were 44 comparisons in which a Cui. or a Con. group was superior to its paired group in conceptual maturity. In 22 of the 44 comparisons, the group rated higher in conceptual maturity surpassed its paired group in proficiency, and in six it was inferior to the latter. In fourteen, the two groups were equally matched in proficiency.

In the English study there were 22 instances in which an H or an L group demonstrated a higher degree of conceptual maturity than did its paired group in the same program (Cui., or Dien., or Con.). In eleven of the 22 comparisons superiority in conceptual maturity was found to be related positively to superiority in proficiency, and in two, negatively. In the single remaining comparison the two groups of the pair showed about the same amounts of proficiency.

Still in the English study: in the three-way comparisons of H groups (and, separately, of L groups) drawn from the different programs, there were fifteen comparisons in which one program group exhibited the greatest degree of conceptual maturity, and in eleven of them the same group was also the most proficient. In no case did the group ranked third in conceptual maturity earn a place higher than second for proficiency.

If one combines the results of the comparisons in the two independent studies, one discovers that superiority in conceptual maturity accompanied superiority in proficiency in 56 of the 81 chances, -- 67%. While these figures do not prove that the higher the level of meaningful arithmetical abstractions, the higher the level of proficiency, they do show that progress in conceptual maturity need not be purchased at the expense of progress in proficiency.

Interpretation

The relative effectiveness of the Cui. and the Con. programs

While it was hardly expected that the Cui. program, found to be much more effective than the Con. program in the Scottish schools, would prove to be less effective than the latter in the English schools, there was always the possibility that this might happen. And it did happen. The fact that it did happen, while it prevents a final, decisive evaluation of the two programs, is by no means disconcerting. On the contrary, among other things, it affords the opportunity to raise and to discuss, later on, certain general problems relating to evaluative research in arithmetic.

But how is one to account for the exchange of the relative positions of the Cui. and the Con. programs in the two independent studies?

As the first step in developing an explanation one must recognize the error in believing that there are such things as "the" Cui. program and "the" Con. program, except in one sense. They actually represent only paper organizations for the guidance of instruction according to differing theoretical frameworks. These theoretic differences are real whether or not they are explicitly stated and elucidated.

The second step is to remind one's self that in school practice paper organizations may take a variety of forms. While the theoretic differences between "the" Cui. and "the" Con. programs persist and influence what is done in the classroom, the quality of instruction according to each of "the" programs may range all the way from excellent to poor.

Now finally, the explanation as the writer sees it: the Cui. program (quotation marks dropped from "the" since the point has been made) was better taught in the Scottish schools than in the English schools, and the Con. program was better taught in the English schools than in the Scottish schools. In a word, in the schools of each country the better taught program, whether Cui. or Con., was the more effective of the two in promoting conceptual maturity.

The evidence to support this judgment is indirect and largely inferential. It was obtained from the following sources: (a) many observations made by the writer when he visited classrooms in English and in Scottish schools some five years ago and again two years ago; (b) conversations with numerous teachers and school officials and with some members of university faculties in Education and other persons who were familiar with changes in the teaching of arithmetic; and (c) the examination of limited amounts of instructional materials. Of necessity, the data in each of these categories are incomplete, for the writer was not bent on making an historical study. Moreover, his impressions and judgments are subject to correction by those better informed, but (he believes) at relatively minor points.

It is of course essential to avoid sweeping generalizations and the uses of black and white only to portray the situation respecting arithmetic instruction as it evolved in the Scottish and in the English schools. To fail to respect these cautions is to exaggerate dissimilarities and to present them as differences in kind rather than as differences in degree, as indeed they were.

With this digression out of the way, let us now look into arithmetic instruction as it has evolved in the last decade or so, first, in the Scottish schools, then in the English schools. Comments will be restricted to the Cui. and the Con. programs since it is with these only that we are now concerned.

In Scotland for some years prior to the making of this investigation some schools had been experimenting with the Cui. program; moreover, experimenting carefully and, if possible, under the direction of one of a number of knowledgeable exponents of the program, some of whom had studied or corresponded with M. Cuisinaire. As a consequence, in such instances teachers who employed the rods tended to learn, both from the practical guidance and from more or less systematic instruction furnished by authorities, what to do with them, how to do it, and why. In a word, in more than a few places the Cui. program was the object of serious study. A goodly number of teachers became expert in teaching the program and, convinced of its special advantages, enthusiastically worked for its more general use. It is not too much to say that there developed something like a reform movement in its behalf, characterized at times by real fervor, --

but fervor grounded in intelligently analyzed experience.

Meanwhile, the Con. program remained practically at dead center. There seem to have been few significant innovations, and, for the most part, teachers in Con. schools continued to teach arithmetic about as they always had taught it, well or not so well.

Now, what was happening in English schools during that time?

The Cui. program did not win converts as quickly and as commonly among English as it did among Scottish school people. There were a few centers, notably in Luton, where this program had been followed for some years, where it was taught apparently with great success, and where it could be seen in operation as it was by many observers. And there were splendidly qualified persons who could, and did, speak with authority and with insight about the program. But they were not so numerous, relatively, as in Scotland; and in spite of their sincere and almost strenuous activities, they were not so influential as were their counterparts in Scotland. Nothing like the Scottish "reform movement" in favor of the Cui. program put in an appearance; and comparatively few closely supervised try-outs were undertaken in the schools.

This is not to say the Cui. program made no headway at all. On the contrary, it was officially "adopted" in not a few school authorities. However, it was by no means rare for boxes of the Cui. rods to be purchased for a school, only to be stored on handy shelves; nor was it rare to select for a try-out a class or two in a school, presided over by teachers relatively untrained in the program and undirected so far as their practices were concerned.*

---

* Of course, these same things happened also in Scottish schools, particularly when a headmaster or headmistress (as in English schools) felt pressured into use of the Cui. program and agreed, very reluctantly, to "give it a try."

Now, one cannot "play around" with the Cui. program or "follow" it in a desultory, unenlightened fashion and still expect it to work miracles. Among other things, expertness on the part of teachers is a prime requisite to success. Otherwise, classroom activities with the Cui. rods may amount to no more than the haphazard manipulation of colored sticks.

True, expertness on the part of teachers is a condition of success in any system of instruction, - in the Con. program as much so as in the Cui. program. But the point is that, in English schools, the Con. teachers as a group already had considerable expertness in their program, whereas the Cui. teachers as a group had yet to acquire it in teaching their program. At least one gained the impression by visiting classrooms that in the Con. schools there was a smaller amount of purposeless pupil activity and a larger amount of planned progress in learning than in (most of) the Cui. schools.

Some may infer from the paragraphs above that, in the writer's opinion, the Cui. program, when well taught, is clearly more effective than is the Con. program, when it too is well taught. If so, let it be said that this inference is incorrect. In the immediately following section an attempt will be made to show that neither of these programs has all the merits of a superlatively good system of instruction and that both seem to be incomplete and/or to have weaknesses of one kind or another.

Apparent weaknesses in the programs

The purpose of the following analyses is to discover within the record of each program the places where it appears to have been least effective in promoting conceptual maturity. The sensible thing to do seems to be to disregard the data for the Cui. program as taught in English schools and the data for the Con. program as taught in Scottish schools, in order to concentrate on the data for these programs when they were better taught. This means that only the results obtained from the Scottish Cui. subjects and from the English Con. subjects (and, perforce, from the English Dien. subjects) will be examined.

The Scottish Cui. program. The comparisons below are based on the assumptions that the addition and the subtraction combinations had been taught with equal emphasis and for the same amount of time and that the multiplication and the division combinations, though probably introduced later, had also been taught with equal emphasis and for the same length of time.

The Cui. program was much more successful in developing abstract and meaningful concepts and grasp of relationships when combinations involved addition rather than subtraction, and multiplication rather than division. (Table 5.) (These judgments are of course based upon the data for pairs of combinations in each operation, and the criterion is average per cent for Uses of Meaningful Habituations.)

The Cui. program was weak also in teaching the rationale of computation in subtraction (Table 8) and in teaching the ability to explain solutions of subtraction and division problems. (Table 10.)

To sum up, the Cui. program was less effective in promoting conceptual maturity when arithmetical tasks involved subtraction and division than when they involved their inverse operations, addition and multiplication, respectively.

The English Dien. program. The assumptions made in the preceding analysis with regard to conditions of instruction on the addition and the subtraction combinations and on the multiplication and division combinations will be made here also, as well as in the next analysis.

The Dien. subjects demonstrated little conceptual maturity in ways of thinking answers for the combinations in any of the operations. (Table 105.)

Meaningful Habituations were employed scarcely at all, and Rational Solutions were used in about one third of the attempts with each pair of interview combinations. There is, therefore, little to be gained by searching for specific weaknesses. Rather, the record of the Dien. subjects reflects lack of concern about early mastery of number combinations, a point made several times before now.

The Dien. subjects were about equally successful in explaining the computations in the addition and subtraction examples, but were more successful in this respect with the multiplication than with the division example. (Table 108.) Again, the results are useless as indicating particular differences in the effectiveness of the Dien. program. Rather, they are better interpreted as the consequences of the special manner in which instruction was scheduled from year to year in that program.

The Dien. program was as effective as was either of the other two English programs in developing ability to give good, intelligible explanations for methods of solving problems. (Table 110.) But the records of the Dien. subjects in this ability are uneven for the sets of problems in the different operations, being much better with addition than with subtraction problems, and substantially better with multiplication than with division problems. The difficulty encountered by Dien. subjects in explaining solutions of subtraction problems points to a real weakness in the Dien. program, but not so their difficulty in explaining division problems, for in the Dien. program work with division is postponed for a considerable period of time.

As was explained in Chapter 2, the Dien. program is a combination of two parts: the part built around use of the M. A. B. and the part supplied locally by teachers under direction. The M. A. B. part should have enabled Dien. children to do exceptionally well in explaining computations. It was none too effective in doing so. The Non-M. A. B. part should have enabled them to do well in verbalizing reasons for solutions of problems; and it was effective at this point.

The English Con. program. To approximately the same extent in each comparison, this Con. program was effective in getting children to use conceptually mature ways of thinking answers for the pairs of combinations in addition and in subtraction, and for the pairs in multiplication and division. The range of per cents for Satisfactory Explanations of the four computations is from 60 for the addition example to 51 for the multiplication example. There is little evidence here of differential effectiveness.

There is, however, evidence that, in enabling children to explain solutions of the sets of problems, the Con. program varied in effectiveness. The Con. subjects were much more successful in accounting for what they did in solving addition problems than in solving subtraction problems, and in solving multiplication than in solving division problems.

Conceptual maturity vs proficiency once again

In the descriptions of the Cui. and the Dien. programs in Chapter 2, statements appear to the effect that, in the first three school years, it is not customary to press for immediate mastery of the number facts and for immediate computational skill. During this period, attention is concentrated, rather, upon the understanding of numbers and of number relationships, as well as of mathematical principles and generalizations of wide application. Supposedly, later on, computational proficiency is stressed, at which time, largely through use, the number facts will be learned.

One may legitimately wonder whether, under these conditions, the number facts will actually be learned as they should be known. Purely by accident, a small amount of evidence on this question became available. In one school authority the writer did not make clear that interviews were to be had with children who were just completing their third year, these and these only. As a consequence, children completing their fourth year were used as subjects, and a new series of interviews had to be held with children who were at the desired point in their schooling.

On the assumption that the younger subjects would, after a year, arrive at the same status in learning as the older subjects had achieved, one can use the records of the two groups to see how delayed emphasis on the mastery of the number facts actually worked out -- in one instance. It is important to note that, in brightness, the two groups were comparable; that they were drawn from the same school, a Cui. school; and that they were interviewed by the same persons. These facts do not, however, make up for the small samples of children involved; hence, the statement above that only "a small amount of evidence" was collected.

The data in Part A of the unnumbered table below illustrate nicely what is meant, in the title of this report, by the phrase, "The Movement Toward Conceptual Maturity": (a) the practical abandonment of thought processes classified under Rubric C (Guessing, etc.) which may or may not represent intermediate stages in the attainment of conceptual maturity; (b) the decline in the use of Rational Solutions, (c) in favor of Uses of Meaningful Habituations. And the data in Part B show that proficiency, however measured, greatly increased, as anticipated, with the use of the number facts in computational practice.*

---

* The data on Proficiency were obtained some time before the data on Thought Processes. This fact probably accounts in some way for the larger per cents of both groups for Non-Attempts in the first encounter with the number combinations, and for the larger per cent of the older group for Uses of Meaningful Habituations than for Correct Answers.

Reactions to the Eight Number Combinations	End of the 3rd School Year (31 Subjects)	End of the 4th School Year (29 Subjects)
A. Thought Processes		
Indeterminate	2%	3%
Non-Attempts	5	0
Counting	4	0
Guessing, Rote Recall, Recitation of Tables, etc.	32	3
Rational Solutions	36	5
Uses of Meaningful Habituations	20	90
B. Proficiency		
Correct Answers	56	84
Non-Attempts	19	5
Prompt Responses	21	75
Long Pauses	40	9

On the other hand (there usually _is_ an "on the other hand"), one is left with further questions not answerable by the data reported. _Why_ should mastery of the number facts be deferred so long, thus requiring children, for a time, to struggle through computations by employing relatively unknown number facts? And _why_ is computational proficiency regarded, too, as something that can properly be postponed? What is the rationale of such an organization of instruction? These questions, and others, lead one inevitably into problems to be discussed in the next section and later on.

Objectives, emphasis, pace of instruction

In Chapter 1 it was explained that for this investigation three kinds of arithmetical task were selected, competence in which seemed to be objectives common to the three programs. Moreover, measures of conceptual maturity were limited to those which could be determined by interviewing children at work with items chosen for each of these tasks: levels of meaningful abstractness in ways of thinking answers for the number combinations, degree of success in accounting mathematically for computations in abstract examples (three of which required the re-naming of numbers), and ability to verbalize reasons for solutions of one-step problems.

But in the British schools the scope of arithmetical content taught in the first three school years extends, in each program, far beyond the simple number combinations, relatively uncomplicated computation, and the solving of one-step problems. For example, the Cui. program provides many experiences relating to fractions. For another example, the Dien. program, through use of the M. A. B., is designed to teach much about the structure of number and about the operation of numeration systems with bases other than 10. How would the Cui., the Dien., and the Con. subjects have compared in degree of understanding and in abstractness of meaningful arithmetical concepts, if the interviews had included items pertinent to these particular types of content, to say nothing of items pertinent to many types not mentioned?

There is no answer for this question. One can but repeat that the selected interview items did not measure -- nor were they designed to measure -- the totality of conceptual maturity inculcated by each and by all of the programs.

Consider now the matter of the relative emphasis put by the three programs on the objectives supposedly common to them all and represented by interview items. The concept of place value and the use of this concept in understanding computation is a case in point. In some Con. programs place value is stressed in the study of the meaning of numbers in the decimal system of notation and prior to its application to computation. In the Cui. program, on the other hand, at least in most of the English Cui. schools, this idea is not taken very seriously in the first three school years. How, then, was it just to such Cui. subjects to expect them to be skillful in using this idea in explaining computation?*

Again there is no answer for the question raised. The relative amount of emphasis put upon common objectives in the different programs would have been difficult to ascertain; and, in any case, it could scarcely have been "controlled."

Consider, last of all, the problem of pace of instruction. As an example, let us take the number combinations once more. In all three English programs, it is safe to say, it is anticipated that children will eventually acquire mastery of the simple number facts. But the critical word here is "eventually." When should children have this mastery? The earlier the date set for mastery, the faster the pace of instruction has to be. Since this date typically comes sooner in the Con. program than in the Cui. and the Dien. programs, was it equally fair to the three groups of subjects to assess level of conceptual maturity at the same point in schooling and according to the same standards of expectation? And would not this same question apply to other parts of the arithmetic curriculum in which pace of instruction differs from program to program?

---

* Explanations of computations making no use of this idea but still mathematically sound were accepted as Satisfactory; but there were exceedingly few instances of this kind.

Still again questions have been asked that cannot be answered satisfactorily. Why, then, ask them at all?

Two purposes are served by asking these unanswerable questions.

(a) At the start of this chapter, the problem undertaken for study in this investigation was said to be far more complicated than it appears to be at first glance. The questions raised above underscore the truth of this statement.

First, in the attempt to evaluate the effectiveness of the three programs by means of interviews, scope of measurement had to be sacrificed to depth, and the evaluation had to be restricted to but three of the many parts of the whole of arithmetical content. Inclusion of more parts, however desirable from the viewpoint of the extent of the evaluation, would have made the inquiry completely unmanageable. After all, interviews with children cannot be indefinitely prolonged. Hence, the relative standing of the programs had to be based upon the responses of children to the three arithmetical tasks alone and can be said to be valid only for these three.

Second, differences among the programs in the relative emphasis given to these three arithmetical tasks had to be neglected. Yet, differences in emphasis are of obvious importance, for they determine the extent of the opportunities afforded for children to advance toward meaningful arithmetical abstractions.

Third, differences in pace of instruction with reference to the three chosen arithmetical tasks had to be dismissed, critical as they are in influencing the degree of conceptual maturity that can be expected at any given time.

So, all in all, the problem set for this study was far from simple. Actually it involved many uncontrolled variables, to which no consideration could be given. In a word, the problem, as it was investigated, had to be over-simplified, in order to get ahead at all with the inquiry.

(b) In view of the analysis preceding (a) above, there is nothing in the findings of the study to distress any reader, be he an advocate of the Cui., or of the Dien., or of the Con. program. If his particular program did not turn out well in the comparisons, he has already been provided sources of comfort. He can say, and with justice: "But the subjects in my program were penalized and never had a fair chance because they could not show how conceptually mature they were in this, and in this, and in this part of our program"; or, "because we prefer to put less emphasis on this objective in favor of more emphasis on that objective"; or, "because we don't hurry our pupils in these and those matters as is done in the other programs, for we leave to time and to experience much of the responsibility for advancing learning."

Implications

Some investigators in reporting their research hold themselves rather rigorously to presenting the facts they have uncovered. They insist that to go beyond the data is not only to engage in a dangerous practice but to be unscientific.

The writer does not belong to this school of thought. It is his view that an investigator, in the course of his study, must gain insights and have hunches which, though they cannot be validated by the data at hand, are nevertheless important enough to be made available to others.

So much the reader must have inferred from the freedom with which the writer has used his judgment -- and has said or implied as much -- in interpreting his data and the findings of comparisons and in attempting to explain such findings. And it is only fair to advise the reader that from here on the writer will be relying upon and stating his judgments even oftener. Actually, were it not for transgressing traditional practice in reporting research, he would write in the first person, the better to indicate his awareness of being in areas of conjecture and of relying upon opinions that are peculiarly his own and, perhaps, shared by no one else.

Improving the arithmetic program in the first three grades

One purpose of this study, not mentioned heretofore, was to see whether there might be found in practices engaged in and in the programs in use in British schools suggestions for improving the teaching of arithmetic in the schools of the United States. The writer believes that this purpose was accomplished, at least to some extent, and then only on the assumption that what is needed are changes in the currently most popular system of instruction in American schools. This is not the place to outline in detail a complete overhauling of this system or to propose and defend the substitution of a radically different system. In the writer's opinion nothing so drastic as either of these alternatives is called for.

The system of arithmetic instruction which for decades has been most extensively used, in some form or other, in American schools resembles the Con. program more than it does either of the other programs studied in the English schools. This is so because counting, rather than experiences with ratio and measurement as in the Cui. program, is used in initial encounters with number. By enumerating sets of discrete objects children acquire the serial idea of number. Next they learn, without enumeration, to recognize collections of discrete objects, in patterns like ⁙ , as units or entities in themselves, and they achieve the group idea of number. Then comes the component idea, learned by taking sets apart and reassembling the subsets, first with objects, then more abstractly; and finally the addition and the subtraction combinations are discovered. Before or after the teaching of the component idea, the meaning of two-digit numbers is taught in terms of place value, thus assuring readiness for the use of this idea in computation. Later is the ratio idea of numbers, and children are introduced to the multiplication and the division combinations. After practice to assure mastery of the number facts, they are put to use in computation and in facilitating skill in problem solving.

This very much over-simplified analysis will be recognized as describing much better the American equivalent of the Con. program as it was taught twenty-five years ago and as it is taught today, for in the intervening years much new content -- arithmetical, algebraic, and geometric -- has been introduced and far greater emphasis is now placed upon mathematical relationships and principles. Nevertheless, the elements of the "traditional" American plan as set forth in the foregoing paragraph are still discernible as undergirding instruction in arithmetic in grades 1 to 3 in a majority of the schools in this country.

In order to make unmistakably clear his position in all that follows, the writer should immediately confess to a prejudice that is implicit in the paragraphs above. He prefers the "traditional" American system of instruction as it has evolved over the years to some of the very dissimilar programs which have emerged during the last ten years. These latter programs have contributed much to the improvement of the "traditional" program; but, in the writer's opinion, they go much too far in such matters as insistence on precision of terminology from the outset and the teaching of mathematical ideas and generalizations long before they have true functional value.

Now, at last, what can be borrowed from the British schools to improve the up-dated "traditional" American system of instruction in grades 1 to 3?

The Con. program. The writer saw nothing in the English Con. schools and nothing in the data for the English Con. subjects to require significant changes in the current form of the "traditional" American system of instruction. Nor did he expect to discover much evidence that the latter should be modified greatly in the matter of objectives, content, and organization. After all, the two programs are very similar in basic theory. On the other hand, he did uncover much to confirm his admitted prejudice.

(a) The American system should continue to regard reasonable competence in computation as an objective to be attained at the end of grade 3, and computation performed abstractly rather than by such concrete aids as the Cui. rods and the M. A. B., which may be useful before then. The English Con. subjects showed that this objective can be realized; and it can be argued that some degree of computational skill is socially desirable even at age ten.

(b) Enumeration and counting should be used in the first stages of learning to deal with quantitative situations, for they would seem to make less severe psychological demands on young children than does measurement. However, they should be supplemented soon in a manner to be mentioned shortly.

(c) By the end of the third grade children should have intelligent mastery of the simple addition and subtraction facts, and, perhaps by the middle of the fourth grade, of the multiplication and division facts. But the adjective "intelligent" is important in qualifying the kind of mastery that should be sought: mastery should be based upon understanding.

There is no necessary reason* to postpone mastery of the number facts, and there are reasons for not doing so. In the first place, thorough knowledge of these facts removes ahead of time a serious obstacle to the acquisition of computational skill. In the second place, intensive study of the number facts provides, in uncomplicated situations, opportunities to discover mathematical relationships and principles of wide applicability.

---

* If other content and other objectives are given priority in the early grades, there is a necessary reason for deferring mastery of the number facts; namely, lack of available time.

(d) Considerable attention needs to be given -- more attention than was evident in English Con. schools -- to the idea of place value in the decimal system of number notation, both as a means of understanding the numbers 10 to 99 (and, later, of still larger numbers) and as a means of rationalizing what is done in computation.

The Dien. program. Although some writers of articles on arithmetic and some authors of instructional materials for this subject seem to have lost sight of the fact, arithmetic in the elementary school has a social aim as well as a mathematical aim. Arithmetic is given a place in the elementary curriculum, not primarily to prepare mathematical scholars (though obviously this purpose is important), but to enable children as children, and later as adults, to deal effectively and intelligently with the innumerable quantitative situations that confront them in the course of daily living.

If the social aim is to be realized, children must have many experiences in using the arithmetic they learn in solving quantitative problems of genuine personal concern to them. It is not enough to give them the usual formalized problems of the textbook (though these can perform a real function), for, to children, these problems are not their problems, except in an artificial, classroom sense. Instead they must come to see that situations requiring the use of arithmetic are common in life and that only by employing arithmetic can these situations be satisfactorily met.

It is not too difficult to supply guidance in teaching the strictly mathematical aspects of arithmetic, for these can be identified and a line of attack can be planned in advance. Such is not the case, however, in providing help in teaching arithmetic as a social study, for the problems which are most likely to arouse children's interest cannot be predicted and programmed. Instead, the teacher must be practically on his own. If he himself is sensitive to the quantitative in life, he will be alert to detect, as they arise, opportunities for children to use arithmetic in ways which are significant to them, - to detect such opportunities and to arrange others. And this is not easy to do.

(a) In providing for the realization of the social aim of arithmetic, the "traditional" American system of instruction, even as it has been modified, is weak, and the Dien. program is strong, at least in the Leicestershire schools. There, under the leadership of Mr. L. G. W. Sealey, Advisor to Junior Schools, a vigorous campaign was initiated some years ago to impress teachers with their responsibility to relate arithmetic to life; and, to judge from the records of the Dien. subjects in problem solving (Chapter 7), the campaign has been very successful. According to Mr. Sealey the attitudes of teachers have been changed, and there are available to assist them files containing, as models or examples, descriptions of the ways in which natural and typical quantitative situations in the school and in the community have been used with profit on previous occasions.

The writer strongly recommends the organization of campaigns like Mr. Sealey's, in order to increase the likelihood that the social aim will be given its proper place in planning instruction. However, it must be emphasized that such efforts must be set up and directed locally; that is, school by school, or at least school district by school district.

(b) Use of the M. A. B., as part of the Dien. program, is another matter. It is the writer's view that the potential values of these ingenious materials are largely lost if they are extensively employed as soon as they are in the Dien. program and for the purposes set for them in that program.* The reasons for this position are both factual and theoretical.

(1) The one place in the interviews where experiences with the M. A. B. had the best opportunity to show their worth was in attempts to explain the mathematics of computation. It will be remembered from Mr. Sealey's statement in Chapter 2 that the M. A. B. are first used with numeration systems with bases other than 10 and that no "special significance is attached to our normal decimal system until the child becomes aware of it for himself." Meanwhile, instruction is designed to lead "a child to an understanding of the processes of addition and subtraction and multiplication and division within a place value numbering system."

Accordingly, the Dien. subjects, who were assumed to have little acquaintance with the decimal system, were not expected to use the words, tens and ones (or units), in their explanations. Yet, they did so, and extensively, as is shown in Table 108. The categories, Words Only and Altered Examples, were employed to classify attempts, both successful and unsuccessful, to explain computations in terms of the place values of the decimal system. On the average, the Dien. subjects used this method of explanation in 83% of their attempts with the four examples. But their per cents for Satisfactory Explanations by this method are consistently and significantly** smaller than are those of the Con. subjects for all four examples and are somewhat smaller or significantly smaller than are those of the Cui. subjects for two examples. Use of the M. A. B., therefore, seems not to have been effective in getting children to use the place values of the decimal system in accounting for what they did in computation, either as the result of direct instruction on this system or as the result of learning transferred from the study of other numeration systems.

---

* In the Dien. schools the M. A. B. are typically introduced in the second year and are stressed in the third year.

**"Significantly" both educationally and statistically. See the closing section of Chapter 3 for the reasons why, statistically, differences between per cents have to amount to at least 10 points to be regarded as "significant."

When children employed methods of explaining computations which revealed no use of the place values of the decimal systems, their attempts were classified as Other Explanations; and when the attempts were successful, credit was given for Satisfactory Explanations. Presumably, Dien. children who employed Other Explanations were drawing upon what they had learned in working with the M. A. B. If so, these experiences were not very helpful in the case of children rated low in brightness and in school achievement.

According to Table 119, 171 of the Dien. subjects were rated 3 or 4 in brightness, and these subjects had a total of 684 opportunities to explain computations. If the per cents in Table 121 are changed to whole numbers, they attempted Other Explanations in 144 examples and earned credit for Satisfactory Explanations in 21 instances, -- in about 1 out of 7 of their trials.

According to Table 126, 184 of the Dien. children were rated 3 and 4 in arithmetic achievement. If the per cents in Table 128 are changed to whole numbers, these children attempted Other Explanations in 132 of their 736 chances, and they were successful 15 times, -- in 1 out of 9 of their trials.*

---

* The 85 Dien. subjects rated 1 in brightness had 340 computations to explain, and they used Other Explanations in 37 instances, and were successful in seven, -- in 1 out of 5 of their trials. Apparently these very bright children (the H group) had profited more from their experiences with the M. A. B. than had the children in the L group. (1) The H group understood better the place values of the decimal system and used them oftener to explain computations, -- in 88% of their attempts compared with 78%. (2) The H group's fraction for successful use of this method (one fifth) is somewhat better than the corresponding fraction of the L group, one seventh.

The samples of L subjects in these two analyses represent 48% (brightness) and 51% (arithmetic achievement) of the total of 358 Dien. subjects. Instructional materials which do so little for half of the school population are open to criticism. However, this criticism need not be directed so much against the materials themselves as against the time when they are used and the manner in which they are used. And it is really at these points that the writer is questioning the value of the M. A. B. in the Dien. program as it is presently organized. He has no doubt whatsoever that, employed later and more slowly, the M. A. B. can contribute greatly to the intelligent learning of arithmetic.

Possibly the two types of evidence adduced (for the total number of Dien. subjects in comparison with the total numbers of Cui. and Con. subjects, and for the Dien. L groups according to two factors) are unfair to the M. A. B., for the criterion employed in judging their worth may be unsuitable, and the M. A. B. may well serve important functions not taken into account. But, to the writer, the evidence as a whole warrants considerable doubt concerning the value of the M. A. B. in the beginning school years.

(2) Dr. Dienes, originator of the M. A. B., bases his system of instruction on four principles*: 1. Dynamic Principle, 2. Constructivity Principles, 3. Mathematical Variability Principle, and 4. Perceptual Variability Principle. Principles 3 and 4, being the most important in the present context, are quoted:

> "3..... Concepts involving variables should be learnt by experiences involving the largest possible number of variables.
>
> "4..... To allow as much scope as possible for individual variations in concept-formation, as well as to induce children to gather the mathematical essence of an abstraction, the same conceptual structure should be presented in the form of as many perceptual equivalents as possible."

---

* Dienes, Z. P. Building Up Mathematics, op. cit; page 44.

Accordingly, Dr. Dienes advocates the use, not only of the M. A. B., but of other perceptual representations of numbers and of number operations; and he believes that true or complete understanding of the decimal system and of its set of place values can emerge only after a succession of prior varied experiences (with the M. A. B. and otherwise) with other numeration systems; hence, the postponement of instruction on the decimal system. One can scarcely doubt that this system will be better understood if other systems also are understood and that experiences with the latter should be provided to children -- but not necessarily before experiences with the decimal system.

It is at this point, the question of timing, that the writer disagrees with Dr. Dienes; but there is no denying the fact that this difference of opinion has large practical consequences in the organization of instruction. The writer wants, as much as does Dr. Dienes, that children should have a truly thorough knowledge of the decimal system; and he believes, with him, that this kind of knowledge cannot be acquired without knowledge of other systems. But to the writer thorough and complete knowledge of the decimal system is not an immediate but an ultimate goal, to be achieved slowly over the years and not in a matter of months, lest the hurried pace of instruction make the goal unattainable.

Certainly it has been demonstrated often enough that, without studying other numeration systems, children can come to have a functional understanding of the decimal system, limited and incomplete as it may be as measured against the ultimate goal. The majority of the English Con. subjects showed that they had this understanding and could use it in explaining computations in terms of place value; and countless American children also have acquired it in the "traditional" program of instruction. Such a degree of understanding should suffice for the time being, to be built upon, extended, and deepened over the

years as other systems are studied through the aid of the M. A. B. or comparable aids.

The Cui. program. There has been a long-standing controversy* with regard to the relative merits of counting and of measurement as means of inducting children into the mysteries of number and of arithmetic. While the theoretical issues have never been firmly and finally resolved, school practice swung steadily, at least in this country, to the use of counting; and it was not until the appearance of such concrete aids and such programs as those devised by M. Cuisinaire and by Gertrude Stern that this movement was checked.

It is now time, the writer believes, to realize that we do not have to choose between a counting and a measurement approach and that there is room, and need, for the better features of each approach to be combined into a single program of instruction. And each approach does have unique advantages, one over the other. The proposed joining of the approaches is underway in British schools. While the Con. program seems to remain relatively unchanged, advocates of the Cui. program are adopting practices that are characteristic of the Con. program. What the writer is suggesting is that the "traditional" American system of instruction, essentially a Con. program, be modified in the direction of the Cui. program. A beginning has already been made through the increasingly popular use of the number line.

Grafting measurement into the American instructional scheme does not necessarily involve adoption of the Cui. rods, though it will require something comparable to them. As others have pointed out, color as used with the Cui. rods has nothing to do with mathematics. And nothing is gained -- and much may be lost -- by having to refer to the orange rod and to the black rod instead of,

---

*** See the reference written almost seventy years ago: McLellan, James A., and Dewey, John. "The Psychology of Numbers and Its Application to Methods of Teaching Arithmetic." New York: D. Appleton, 1896. 309 p.**

more precisely, to the 10-rod and the 7-rod, respectively. Such quantitative names cannot be employed at the outset because enumeration is disapproved and because the Cui. rods are not proportionally segmented, or segmented at all, for that matter.

Of the many values of experiences with measurement only a few need be mentioned here.

By enumerating objects in a set children begin to acquire the idea of natural numbers. They learn that the natural numbers start with 1, followed by 2 (1 + 1), then by 3 (2 + 1), and so on. In a word, the next natural number after n is always n + 1; and the natural numbers are "ordered," each one after 1 being larger by 1 than is its immediate predecessor.

The natural numbers, as "counting numbers," can be used to enumerate the number of rodents in a cage, of animals in a picture, of pieces of furniture in a room. And they are appropriate in these situations regardless of the relative size, the shape, the color, and the spatial distributions of the objects in the set. Although the natural numbers themselves are "ordered," the objects enumerated may be taken in any order.

The natural numbers as used in enumeration are similar to the natural numbers as used in measurement, in that the same names and the same order are employed in both. However, they are also dissimilar in the two instances. To use the number line for purposes of illustration: the points on the number line designated by successive natural numbers are equi-spaced. That is, the distance between any two successive points assigned the names of natural numbers is equal to the distance between any other two successive points bearing the names of natural numbers.

Hence, activities with the number line, or with other aids comparably designed, are more useful than are activities involving counting and enumeration alone, in extending children's understanding of numbers. For example, with such aids they can discover that there is no natural number corresponding to (4 - 7) and that the answer can be designated only by using a different kind of number name.

To limit the arithmetic program in the early grades to the use of counting is to increase greatly and needlessly the difficulty of demonstrating the concept of fractions. For example, in enumerating the number of children on the beach one is unlikely to encounter half a child. On the other hand, study of the number line or any similar measurement device reveals that between any two points designated by natural numbers there is an indefinitely large number of other points which can be named as fractions.

As a last example: measurement is superior to counting and enumeration as a means of introducing and illustrating the idea of ratio. Take the case, How many 3's are there in 12? Laying out rods or sticks or plastic strips to represent 12, the child next "measures" 12 by putting end-to-end enough 3-rods or 3-strips to equal the length of the representation of 12. On the other hand, in a counting approach with discrete objects, the child may first make a set of 12 marks, then enumerate and encircle a subset of 3, and another, and another, and still another, and finally enumerate the set of encircled subsets to get his answer. The latter approach is relatively more clumsy, time-consuming, and open to error than is the measurement approach.

Evaluative research in arithmetic

This investigation is the writer's first -- and probably his last -- venture into evaluative research. Happily for him, he has had some of his biasses confirmed; but he has also learned several lessons, all of them of similar import: Evaluative research is extraordinarily difficult to plan and to execute if it is to be anything other than routine and superficial in character.

(1) It is relatively easy to obtain quantitative data; but in themselves these data say next to nothing. To make them meaningful and significant, data must be interpreted. Hence, the role of judgment is greatly magnified, enough so as to reduce materially the hoped-for objectivity of the undertaking.

(2) The investigator is in trouble when he has to decide upon the kind or kinds of data he will procure. If he administers an original or a standard test of the typical sort, he may be able to learn how accurately and how quickly children can compute and can solve formal problems, and he will have measures of proficiency only, without knowing how and why his subjects do what they do. That is, he will have no measures of process, to reveal to him whether children perform intelligently or mechanically, with a minimum of understanding.

If he chooses to collect process measures, he must be prepared, together with other persons whom he has trained, to spend untold hours in observing and questioning children individually. So, he is tempted to collect proficiency measures alone. To do so is to disregard the fact that it may be less important to evaluate programs for their relative effectiveness in developing skill than for their relative effectiveness in developing intelligence in dealing with quantitative matters.

(3) The investigator is always forced to limit the scope of his study; for example, in the content to be included and in the objectives with which he will be concerned. He thereby restricts the extent to which his findings can safely be generalized. And the more penetrating his study -- the deeper he goes into his problem -- the more he must sacrifice scope.

(4) To intimidate the potential research evaluator a bit more: arithmetic programs differ in their objectives. For instance, in some programs the social aim of arithmetic is taken seriously; in others it does not figure prominently,

if at all. Shall the investigator collect data relevant to this aim for all programs regardless of the degree to which they stress it? If he does not, why not? If he does and if he makes the social aim his evaluative criterion, what kinds of relevant data can he obtain; and will he not be prejudging the case against programs, otherwise excellent, in which the social aim is viewed as relatively unimportant? If, on the other hand, the social aim is but one of several evaluative criteria, what weight will he attach to it in comparison with the others? And how can he justify his weighting?

Again, programs differ in the emphasis given mathematical understanding and insight as an objective of instruction. Some programs assure only as much as is absolutely needful at the time; others go the limit, extending such understanding and insight to a point deemed suitable for children who hope to become professional mathematicians. Shall the investigator turn a blind eye to these differences with respect to the same objective; or shall he try to do something about them? If the latter, does he not face problems like those mentioned in the preceding paragraph?

(5) In another connection, it has been pointed out that programs are unlike each other in setting the time when objectives that are common to them all are to be realized and when subject matter, again common to them all, is to be mastered. If a late date is set, the pace of instruction can be comparatively slow. If an early date is set, the pace must be comparatively fast. Since he cannot control such differences in instructional schedules save by altering programs themselves, how is the investigator to take into account the inevitable differences in the amounts of learning which programs will have produced at any given time?

(6) The success of any program is very largely dependent upon what is done with it. A program, admirable in its theoretical framework and in its organization,

can be ruined by poor teaching. Conversely, another program, not nearly as commendable, can be made to appear to be very effective if it is well taught. Hence, the investigator may, in the end, be evaluating, not programs of instruction, but quality of teaching. Such proved to be the case in the comparisons of the Cui. and the Con. programs in this study. How is the investigator to prevent this from happening in his research? How is he to assure that all programs he is evaluating are taught equally well?

The list of problems which confront the investigator who seeks to evaluate arithmetic programs could be extended indefinitely, -- in one way, by going into technical questions concerning procedures; e.g., the conduct of the interview and the statistical treatment to be accorded the data. The writer, however, prefers to mention problems which are less commonly considered or even recognized. And he stops his list with six, for these six are enough to give "the miseries" to almost anyone who is about to organize a project in evaluative research. What can he do about these problems? Quite immodestly, the writer ventures to offer a few suggestions.

In the first place, these problems present difficulty largely in proportion to the scope of the proposed research. In his opinion it is impossible, by research, to evaluate programs in their entirety,* so as to include the full range of arithmetical content taught and the full range of accepted objectives. (Least of all is there any justification for claiming over-all evaluation when only parts of the content and only a small number of objectives are represented in the inquiry.) Of course, the problems cited above are present in evaluative research that is not so ambitious, but perhaps less critically, for there may be ways of dealing with them.

---

* If the writer is correct in his opinion, investigations which undertake evaluation international in its scope are hopeless from the outset. How is one to deal satisfactorily with the issues involved in the problems cited above, to say nothing of differences among countries in philosophy of education and in conceptions of the function of arithmetic in the curriculum? Can such investigations do more than to show how children perform on the particular instrument administered, which can scarcely be equally fair and comprehensive in as far as arithmetic instruction in different countries is concerned? That is, can they furnish anything but accumulations of data signifying little until others have completed the more difficult task of interpretation, with all its perils? Tables of test scores do not "evaluate." They provide only the means, or part of the means, to "evaluate."

In the light of these remarks, the investigator is advised to limit his inquiry to reasonably manageable dimensions and to focus his attention upon selected content and selected objectives. But he needs to remember always that he has restricted his study in this manner and to write whatever he does write in full recognition of this fact.

In the second place, in order to deal with differences among programs with reference to content and objectives more effectively than is customary, it is suggested that the investigator (some investigator) try a strategy which, if not new, is certainly novel. This strategy requires two series of tests or interviews or observations. For simplicity let it be assumed that the procedure selected is that of testing and that the testing is restricted to content alone. (The proposed plan should be equally useful -- or useless -- if the testing had to do with objectives instead.)

One testing would involve only content common to all programs. The other testing would involve differences in content, and the instrument would include items for content peculiar to each program. But this instrument would be used with all the children in all the programs. For children in each program, the test would provide both additional measures of achievement (on content taught) and measures of ability to transfer learning (to untaught content). It should not be too troublesome to separate the two sets of scores on the instrument and to treat them accordingly.

If feasible, this plan would insure rather comprehensive coverage of the content in each program, with injustice to none. And for this reason and others, it would broaden the basis of evaluation. But the plan might prove to be wholly impracticable; for example, the testing required might require too much school time. And if the plan should introduce no serious obstacles in so far as the collection of data is concerned, it would certainly not lessen but would probably greatly increase difficulty at one foreseeable point; namely, the determination of the relative weights to be assigned for success with taught and with untaught content.*

In the third place, the investigator would do well to realize the responsibilities he assumes when, as he must, he interprets and attempts to explain his findings. In preparation he should ground himself thoroughly in the area(s) of his study so that he can speak with confidence and with a degree of expertness that will command the confidence of others. He should frankly admit any prejudices he has, carefully preventing their dictation of the conclusions he draws. He should recognize and make a point of stating the assumptions and hypotheses that direct his planning of his attack upon his problem, step by step. He should never disguise the fact that many times, in the absence of supporting data, he must enter the territory of conjecture and rely upon his own judgment. And he should cite alternative interpretations and explanations that may be as good as, or better than his own.

These last suggestions to the investigator are important in all research, but are especially important in evaluative research, which quickly leaves the solid ground of reporting facts for the uncertain terrain of judgment. Here the investigator's equipment consists of the knowledge he has acquired and the opinions and hunches he has formed.

---

* It should be obvious that this system of teaching, as extensive as it is and as voluminous as might be the data procured, does not constitute an evaluation of the selected programs as wholes. For one reason, too many aspects of the programs will have been neglected. For another, the collecting of data is only the initial stage in evaluation, for the data must undergo the intricate and debatable process of interpretation. Hence, the writer, in suggesting the proposed strategy, is not inconsistent with the position he has stated earlier, to the effect that "it is impossible, through research, to evaluate programs in their entirety."

Related Research

It is most unusual for a report of this kind to contain no bibliography, not a single citation to related research, and only a few references to discursive and theoretical articles on ideas and issues similar to those encountered in the making of this investigation.

The reason for the omissions is certainly not a dearth of printed matter. On the contrary, there is a vast body of literature that could be drawn upon, some of it of a research nature, some of it theoretical, or descriptive, or expository. Indeed, it would be more accurate to say that there are vast bodies of literature, -- for example, on the interview and observation as research techniques, on concept formation, on the statistical treatment of data, on principles of curriculum construction and use, on the teaching of arithmetic, on evaluation and evaluative procedures, and so on and on.

Why then the neglect of this literature, or of these literatures?

To start with the research literature: one needs not read far in appropriate sources to find numerous reports of evaluative research, some of them in the field of arithmetic. As a matter of fact there are available published reports of several recent American studies and of at least two British studies which were designed to determine the relative effectivenes of arithmetic programs, one of them the Cui. program. In all of these studies paper-and-pencil tests were used, and in all except one the data procured were restricted to measures of skill in computation and problem solving. In none was any effort made to look into the thought processes of children and to determine relative progress toward conceptual maturity under differing systems of instruction. Hence, it seems pointless to cite and to review them here.

As for related research: that on concept formation would seem to be the most valuable in this context; but it turns out not to be very helpful. Concept formation, after decades of relative neglect in America, if not in Europe, is now fortunately a popular object of psychological research. However, from the viewpoint of educational practice, little of what is done bears even remotely on problems of classroom teaching and learning. Performed commonly under laboratory conditions, with individual subjects, and on learning tasks seemingly resembling not at all those set in the school, psychological research appears to approach concept formation in an extremely artificial and academic manner.

These statements imply no criticism whatsoever of the psychologist who studies concept formation. After all, everyone has the right to investigate whatever interests him in the manner which best suits his fancy. More important, however, is the point that the psychological researcher does what he does for entirely legitimate purposes, to advance his science through the formulation of basic theory. It just happens that his purposes do not coincide with those of persons whose immediate goal is the improvement of instruction, and there is no reason why they should have to coincide.

To turn now to the non-research literature: why has it been neglected? The truth is that it has been neglected only in one sense, namely, by lack of citations to specific references. Over the long years the writer has browsed through or read carefully large segments of this literature and has absorbed from it a host of helpful ideas. The word "absorbed" is used advisedly, for these borrowed ideas have become so much a part of his own thinking that he has no memory of their sources. It would be futile for him to try to give credit where it is due, -- futile and unfair, for in the process of selection he would certainly do injustice to many persons to whom he is deeply indebted. It seemed preferable, therefore, not to undertake a listing of non-research references whose contents are related in one way or another to some aspect of this study.

So much for the absence of the customary bibliographical matter in this report. The writer doubts that it will be missed, for, without it, the reader has had plenty to occupy his attention.

APPENDIX

	Page
Interview Problems	**343**
English Study. Reading as a Factor. Tables 133 to 139	**344**

INTERVIEW PROBLEMS

Skeleton form

Directions

1. Present informally - orally.
2. Assure understanding of each problem setting.
3. Provide no clues to mathemtical operation.
4. Be sure to emphasize underscored word(s) in each problem.

1. Once there was a boy called Sam. He had 16 eggs in a basket. 7 of them were cracked. How many eggs were not cracked?
2. We use 5 sheets of paper to make each of our books. How many sheets of paper will we need to make 5 books?
3. When John had a birthday he had a party. His mother had to get 5 chairs for visitors and then 6 more chairs for other visitors. How many visitors had she altogether?
4. There are 16 biscuits in a box. 4 boys are going to share these biscuits equally. How many biscuits will each boy get?
5. John had 14 fish in a big tank. He had 9 fish in a small tank. How many more fish are in the big tank than in the small tank?
6. One day John's class had to do a sum test. John got 9 sums right and 7 sums wrong. How many sums had he done?
7. 20 children of our class take school dinners. 4 children sit at each table. How many tables are needed for our class?
8. A bag of marbles was shared equally between 3 boys and each boy got 4 marbles. How many marbles were there in the bag to begin with?
9. Susan had a box of 12 crayons. She can find only 8 of them now. How many of her 12 crayons are missing?
10. Tom had 18 books to put away in the cupboard. He put them in equal piles of 6 books each. How many piles of books did he make?
11. John saw 5 boxes of eggs in the shop. Each box had 6 eggs in it. How many eggs were there altogether?
12. Tom and his friends bought some apples. They ate 5 apples, and 5 apples were left. How many apples did they buy?

Table 133. Reading Achievement as a Factor.
Composition of the H and L Groups in Each Program

Bases of Comparison	Cuisinaire		Dienes		Conventional	
	H	L	H	L	H	L
Number of Subjects	75	110	108	144	127	108
Per Cents by Sex						
Male	36	64	44	56	50	59
Female	57	43	56	44	50	41
Weighted Means for						
Total Instructional Time	558	586	563	615	616	578
Age	7 yr. 9 mo.	7 yr. 11 mo.	7 yr. 11 mo.	8 yr. 1 mo.	8 yr. 0 mo.	7 yr. 11 mo.
Estimated Ratings:						
Brightness	1.13	2.90	1.45	3.22	1.36	3.13
General Achievement in Arithmetic	1.00	3.29	1.00	3.42	1.00	3.39

Table 134. The Number Combinations. Thought Processes with the Eight Combinations; Averages in Per Cents

Rubrics and Categories	Cuisinaire		Dienes		Conventional	
	H	L	H	L	H	L
A. Non-Attempts	4	15	8	18	2	5
B. Counting	12	30	24	38	17	27
C. Guessing	2	7	2	7	3	10
Rote Recall	5	4	7	3	2	1
Recitation of Tables	11	8	8	6	10	11
Mechanical Uses: Reverses and Inverses	4	7	6	2	2	5
Sub-Totals	22	26	23	18	17	27
D. Rational Uses:						
Reverses	1	0	0	0	1	1
Inverses	2	1	3	0	3	2
Solutions, Other Forms*	45	24	34	21	28	21
Sub-Totals	48	25	37	21	32	24
E. Uses of Meaningful Habituations	11	2	6	2	30	15

* The category, Inverses, Plus, is included

Table 135. Computation. Quality of Explanations of Computations, Attempts Only; Averages in Per Cents

Classification of Explanations*	Cuisinaire		Dienes		Conventional	
	H	L	H	L	H	L
Satisfactory						
Words Only	13	5	28	13	26	14
Altered Examples	33	19	20	14	36	29
Other	0	0	3	1	0	0
Unsatisfactory						
Words Only	34	43	30	42	18	29
Altered Examples	19	25	10	7	19	25
Other	0	8	9	21	1	2

* The category Imagery is omitted

Table 136. Problem Solving. Reasons Given for Solutions, Attempts Only; Averages in Per Cents

Quality of Explanations	Cuisinaire		Dienes		Conventional	
	H	L	H	L	H	L
Good	62	38	58	45	60	43
Unsatisfactory	14	25	9	27	10	18
Indeterminate	3	4	3	2	6	10

* The category Fair is omitted

Table 137. The Number Combinations. Proficiency in Giving Answers: Accuracy and Rate; Averages in Per Cents*

Nature of the Measures	Cuisinaire		Dienes		Conventional	
	H	L	H	L	H	L
Accuracy						
Correct Answers	85	63	80	56	87	70
Non-Attempts	4	16	10	19	2	6
Rate, Attempts Only						
Prompt Responses	50	19	46	21	53	29
Long Pauses	11	35	17	40	11	29

* Per cents omitted for the categories, Incorrect Answers and Slight Pauses

Table 138. Computation. Proficiency: Accuracy and Quality of Performance; Averages in Per Cents*

Types of Measure	Cuisinaire		Dienes		Conventional	
	H	L	H	L	H	L
Accuracy						
Correct Answers	71	47	60	32	84	62
Non-Attempts	4	20	18	36	1	6
Quality of Performance, Attempts Only						
Smooth	83	69	65	35	85	69
Erratic	16	32	35	64	15	30

* Per Cents omitted for the category, Incorrect Answers

Table 139. Problem Solving. Proficiency: Accuracy and Performance Level; Averages in Per Cents*

Types of Measure	Cuisinaire		Dienes		Conventional	
	H	L	H	L	H	L
Accuracy						
Correct Answers	87	69	76	72	90	76
Non-Attempts	2	10	4	7	2	4
Level of Performance, Attempts Only						
High	53	28	45	27	68	51
Low	12	44	14	36	12	22

* Per cents omitted for the categories, Incorrect Answers and Intermediate Level of Performance